The ecological context of diseases of public health importance

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Bethany Anne Levick

May 2017

Preface

I hereby declare that this thesis consists of my own work, unless otherwise stated, and has not previously been submitted to the University of Liverpool or any other University in application of a higher degree. This text does not exceed 100,000 words and meets the formatting requirements of the University of Liverpool.

Contents

Abstract

The ecological context of diseases of public health importance
Bethany Levick

The prevention and control of infectious disease continues to be a central priority for human public health. Two infectious agents, the bacterium *Yersinia pestis* and the filarial nematode *Onchocerca volvulus* have been responsible for significant public health burden in many populations, and challenges remain in the control and prevention of them both. With known environmental and climatic drivers, vector based transmission and in the case of *Y. pestis*, zoonotic reservoirs, fully characterising the ecological context of both diseases is vital to properly understanding their epidemiology. In the present study, ecological field data is utilised to investigate this, towards improving predictions and designing control strategies. Despite being contrasting infectious agents, the role of spatial scale and resolution are identified to be vital aspects of both investigations. Ultimately, the importance of a holistic approach to infectious disease epidemiology is highlighted across both systems.

In humans and some other animals, infection with *Yersina pestis* causes the disease plague. Although human infections are now limited to the hundreds annually, they disproportionately affect individuals in lower income settings and as such are still of public health concern. A number of wild rodent species tolerant to infection with *Y. pestis* act as reservoirs for the bacterium, from which it can transmit into less tolerant susceptible populations. To this end, much research has focused on predicting periods of high infection prevalence (epizootic outbreaks) in the reservoir species in order to predict and prevent transmission events into less tolerant populations.

One such reservoir species is the great gerbil, *Rhombomys opimus*, residing across the pre Balkhash desert in Kazakhstan. *Y. pestis* prevalence in this population is known to be strongly associated with the gerbil population size. The present study aims to resolve current ability to predict the gerbil population. Burrow level environmental properties are identified to be predictors of gerbil presence at, and long term occupancy of burrows. Gerbil occupancy of burrows could not be replicated using standard metapopulation approaches, nor related to the connectivity of the burrow.

Atypical strains of *Y. pestis* lacking the fraction 1 (F1) surface antigen are isolated and found to be common in the gerbil population. This antigen is highly immunogenic and has a role in virulence determinance. However, no significant hetereogeneities in their distribution or associations with epidemiological outputs could be identified.

In several African countries with endemic levels of infection with *Onchocerca volvulus*, an unusually high prevalence of epilepsy has been observed in the population. Here cross sectional data from a population in the Democratic Republic of Congo is used to build a picture of epilepsy prevalence and to explore a possible functional relationship between epilepsy and *O. volvulus*. The relationship between *O. volvulus* and epilepsy is strengthened, and spatial and ecological data lead to the suggestion that responsible drivers have local, rather than global relationships with epilepsy.

Acknowledgements

I am fortunate to be surrounded by a great many wonderful people who have provided all manner of support, advice and encouragement over the past three or so years. I can only name some of them here, but rest assured that all the individuals I have had the pleasure to share time have made this a hugely rewarding experience, if often equally a great challenge.

First, I would extend an immeasurable gratitude to my supervisors, Prof Mike Begon and Dr Stephen Cornell. Stephen, in particular I want to thank you for your encouragement for me to extend and challenge my technical ability, and building my confidence to do so. Mike, your advice and support has aided me in so many ways over the past few years. I will carry your professional and scientific guidance throughout my career, and hope to maintain the high standards I am now used to being held to.

Dr Anne Laudisoit, thank you for our countless discussions and sharing your great wealth of knowledge and insight on both the projects in this thesis. As a quantitative researcher, I am hugely grateful to the field teams without whom my work wouldn't be possible: especially Aidyn Yeszhanov, Yermek Idkazbaimbetova & the trapping teams in Kazakhstan, and all of the field teams and participants in the DRC.

I have made numerous excellent friends from across the Institute of Integrative Biology, many from the Department of Ecology, Evolution and Behaviour. We have all made a great many fun memories and grown together. I wish everyone the absolute best for whatever your future holds.

I cannot thank my fiancé Liam enough times for his undying patience, love and care. I am sure you're looking forward to this being done nearly as much as I am looking forward to us finally getting married in November.

I am relatively certain that my parents' offer to support & help me "whilst ever I remained in education" was made with absolutely no anticipation that I would do so until the age of 25! Even so, they and my family have been a consistent light of positivity and care throughout my educational career.

Cheers everyone, I've had a good 'un.

"I don't give up. Never have. Never will."

Jonah Lomu, in an interview with *The Guardian* newspaper, 2010.

Author contribution

This thesis uses several databases collected from relevant field sites. Throughout, details of the methods used for field data collection are provided to provide context to the investigation and associated data. It should be noted however that the author was not involved in the design of these field studies or the collection of data. The author's work comprises the design, implementation and interpretation of the data analysis and underlying research questions, and (where relevant) simulation studies described here, and the biological interpretation of the findings in the context of the disease systems.

List of tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of figures

Chapter 1

Chapter 3

1. Ecological field data provides insights into disease ecosystems with implications for public health.

Infectious disease in humans & animals does not occur in isolation, rather representing a mark left by interactions with the ecological community, and ultimately the wider ecosystem within which species act. Public health guidance to minimise exposure risk often does not incur large scale collection or analysis of ecological data. However, data on these aspects is key to understanding the wider context of infectious disease, particularly zoonoses.

Here two large bodies of field data are utilised to quantify ecological processes relating to the transmission of two infectious diseases affecting human populations. *Yersinia pestis*, a bacterium known to cause severe disease in humans and other mammals persists largely in tolerant reservoir hosts. One such reservoir host, the great gerbil *Rhombomys opimus* has demonstrated clear relationships between its population dynamics and disease dynamics. Here unresolved questions about the nature of these associations are investigated, towards an aim of improving predictions of large outbreaks of infection in the animals, an indicator of risk to human populations that may contact them.

Second, reports of epilepsy, particularly of a recently identified atypical form known as nodding syndrome (NS) are unusually frequent in several countries in Africa. Although a specific aetiology for this remains to be identified, high prevalence of epilepsy is frequently found to be concurrent with onchocerciasis (*Onchcocerca volvulus* infection) endemicity. Through a combination of a large cross sectional study of humans, reports of their exposure to ecological niches of the vector (*Simulium spp.*), and environmental data of the surrounding area, potential ecological drivers for the manifestation of epilepsy in association with onchocerciasis are investigated.

1.1 *Yersinia pestis* epizootics in a community of wild rodents in Kazakhstan.

Yersinia pestis is a gram-negative bacterium, capable of infecting humans and a number of animal species, but typically residing in communities of tolerant wildlife species. These are therefore referred to as reservoirs for the bacterium. In humans, infection with the bacterium manifests as the disease plague. Plague is fatal if

untreated, but can be resolved with antibiotic treatment, generally streptomycin or tetracycline based (Perry & Fetherston 1997). Whilst vaccines do exist, their efficacy is still being improved (Derbise et al. 2015). Plague is a disease of great historical and cultural relevance, as the cause of famous outbreaks including the Justinian Plague in the 6th century and the Black Death across Europe in the 13th century (Stenseth et al. 2008). Whilst human infections annually now total under the hundreds, and deaths around one hundred a year, the disease is far from eradicated and particularly in lower and middle income countries remains of public health importance (Bertherat, 2016).

Whilst causing the most harm, *Y. pestis* infections in human populations form a small part of the overall epidemiological cycle of the bacterium. A number of wild mammal populations, generally rodents, are tolerant to *Y. pestis* infection and can sustain endemic infection over long periods. Infection in non-tolerant populations then arises following spillover from these reservoirs (Stenseth et al. 2008). Such tolerant populations exist across Asia, Africa and the Americas, but are restricted geographically and ecologically to "foci" of infection (Figure 1.1). Each focus has its own epidemiology, determined by a mix of biotic and abiotic factors affecting the system. These wild reservoirs make up the sylvatic part of the plague epidemiological cycle and are responsible for maintaining isolates of the bacterium over long periods. A shared pattern across these systems is the apparent re-establishment of *Y. pestis* prevalence into epizootic events, following periods of very low (typically non detectable) prevalence (Gage & Kosoy 2005; Eisen & Gage 2009). The mechanism driving this behaviour remains unknown.

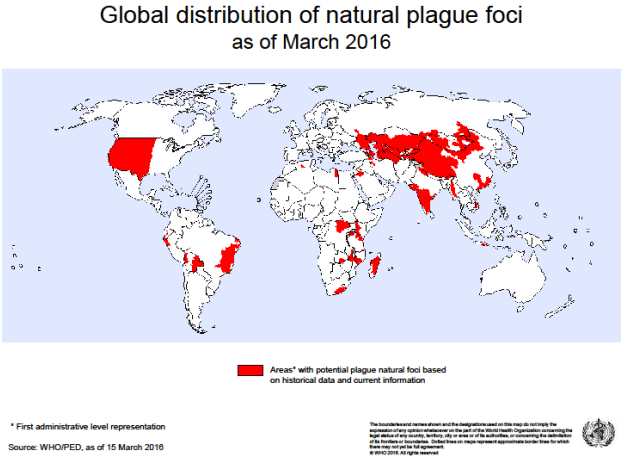


Figure 1.1. Global distribution of Y. pestis foci (WHO 2016).

Rodent species living alongside humans, including the Norway Rat (*Rattus norwegicus*) and the black rat (*Rattus rattus*) are referred to as "commensal rodents", (in reference to their relationship with human populations). These commensal species (typically *R. rattus*) can acquire *Y. pestis* infection via fleas that have previously fed on infected wild rodents. Once infecting species living in close proximity to humans, the bacterium can then move into infected human populations through flea bites. This generally occurs as numbers of the commensal species start to fall due to deaths. Once in human populations, the bacterium can transmit between them and the commensal rodents via flea bites, or eventually by droplets between humans (Stenseth et al. 2008, Figure 1.2). Recently this paradigm has been challenged, and a more diverse range of epidemiological pathways highlighted for *Y. pestis*, involving a larger variety of species and more directly incorporating environmental and ecological drivers (Zeppelini et al. 2016). Whilst not a focus of the present study, the role of climactic fluctuations affecting *Y. pestis* dynamics is well documented, particularly so for the Kazakh system (Schmid et al. 2015; Kausrud et al. 2007; Kausrud et al. 2010). This serves only to reinforce the importance of understanding the dynamics of the wild systems, and to utilise the power they represent as an epidemiological tool.

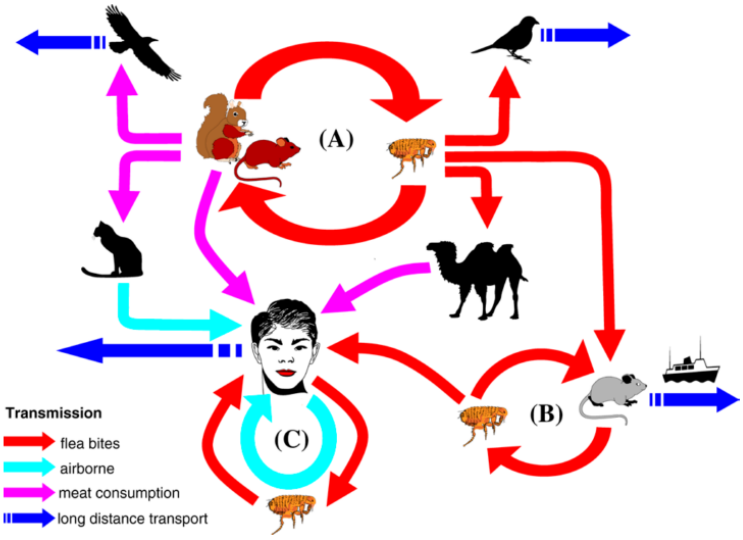


Figure 1.2. The *Y. pestis* epidemiological cycle (Stenseth et al. 2008).

From the above, it is clear that the presence of the infection in non tolerant populations is dependent upon the dynamics of infection in the tolerant wild species. As such, much modern plague epidemiology is anchored in establishing and informing quantitative methods (statistical and mechanistic models) for describing and

importantly anticipating epizootic events. Examples of this are seen across the different plague systems, recent examples include the role of droughts on epizootic events within the US foci of *Cynomys ludovicianus* (prairie dogs) (Eads et al. 2016) and genotypic characterisation of strains present in Madagascan *Rattus rattus* (black rat) infections (Riehm et al. 2015). The data here concern the Kazakh plague focus.

Y. pestis is endemic in a population of wild great gerbils (*Rhombomys opimus*), found across the steppe region of Kazakhstan in the areas surrounding Lake Balkhash. The semi arid desert region covers a 200 x 250km area, and encompasses a variety of environments, vegetation and geography, and features a number of human settlements (Wilschut et al. 2013). *R. opimus* are a fossorial rodent species, with a social structure unique amongst related animals (Randall et al. 2005). Each burrow will house a family unit typically comprising of an adult male, an adult female and several juvenile animals. Some of the young, typically males, will leave the original burrow group to establish new groups (Randall et al. 2005; Begon et al. 2006; Naumov & Lobachev 1975). These may be in existing, but currently unoccupied burrows, or less often animals may construct new burrows. This process is however necessary, as burrows can become damaged beyond use. The burrows typically cover a roughly circular area with a radius of 30m, with access holes on the surface leading to a complex system of underground tunnels. These burrows have specific structures, with dedicated sections serving as food storage and living areas (Naumov & Lobachev 1975).

The *R. opimus* population and their prevalence with *Y. pestis* have been studied in the region since the end of the second world war. Originally this was largely for public health monitoring, to guide control strategies (typically culling and insecticide spraying) to prevent spillover events. These data sets are unique in their temporal continuity and geographical breadth in describing the numbers and infection levels of a wild population.

Research from European groups began in the area in the early 2000's. Using the original Kazakh public health records, and new additional field records generated with closer resolution approaches, four key findings were identified. First, it was identified that population cycles in the gerbils were reflected in epizootic cycles of infection prevalence (Davis et al. 2004). This represents a powerful tool epidemiologically as the current density of gerbils was shown to be predictive of the

future infection levels. Moreover, this remains one of few examples of empirical evidence for threshold effects in disease outbreaks amongst wild species. The concept of a critical community size is well documented in human epidemiology, most clearly demonstrated in the case of measles infections in children in the UK (Bolker & Grenfell 1995). However, even at present, whether a parallel phenomenon exists in wild populations remains a point of debate. A central problem for this lies in the difficulty of collecting sufficient and appropriate data from wild populations, especially concerning transmission networks and fully characterising periods of disease absence (Lloyd-Smith et al. 2005). The example from the Kazakhstan system remains one of the most compelling and complete evidence of such an effect, although it is not completely understood.

Capture mark recapture studies and metapopulation analysis were able to demonstrate that within burrow transmission has no impact on estimates of epizootic spread across the population (Davis et al. 2007; Begon et al. 2006). As such here, and in associated literature, the gerbil population is described in terms relating to the number or density of occupied burrows, rather than the number of individuals. Similarly, infection statuses are ascribed to a burrow group.

Subsequently it was identified that the movement of *Y. pestis* infection across the gerbil population can be well modelled as a percolation phenomenon (Davis et al. 2008). Considering a central, infected burrow across an area of burrows of which some harbour susceptible individuals and some are empty, the spread of infection from the central burrow is restricted by the number of occupied burrows surrounding it. A higher density of occupied burrows creates more connections across the landscape, allowing the infection to percolate outwards (Figure 1.3).

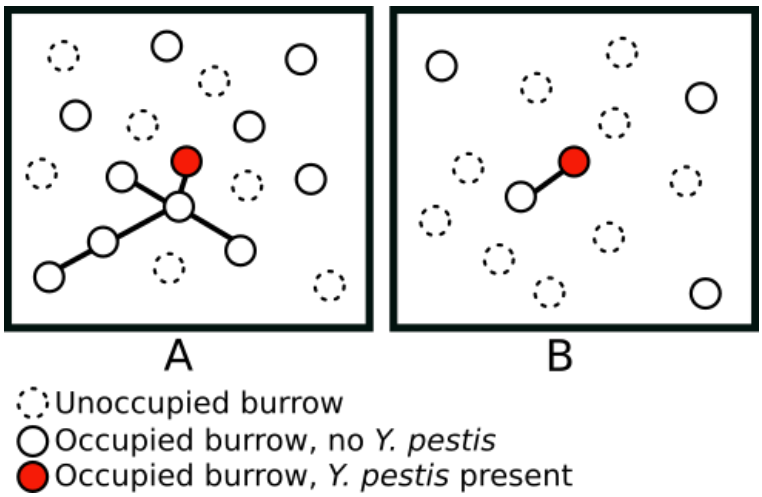


Figure 1.3. Demonstrating percolation of infection. In panel (A), the infection in the central, red burrow is able to spread outwards, via connections provided by the occupied (solid outline) burrows. In (B) there are insufficient connections to occupied burrows available, and the disease cannot spread through the unoccupied (dashed outline) burrows.

The fourth key finding is that the threshold for *Y. pestis* outbreaks is best modelled not only as a single threshold of gerbil density, but a curved threshold describing gerbil density and flea burden amongst the gerbils (Reijniers 2012). The shape of the threshold is further altered by temporal considerations (Reijniers 2014).

The current investigation is the most recent body of work towards identifying possible reasons for the discrepancy between predictions from the curved threshold model and the observed epidemiological data. Specifically, burrows are regularly identified in areas with sufficient gerbil density and flea burden to fall above the threshold (suggesting they should have *Y. pestis* infection), but where *Y. pestis* is not observed. However, all of the burrows below the threshold are not observed to have *Y. pestis* infection, and all of those with *Y. pestis* infection are above the threshold (Figure 1.4). The model is then considered to "overpredict", which whilst preferable from a conservative approach to disease control, is not ideal for increasing efficiency of control efforts. More fundamentally, it suggests that there is an important aspect of *Y. pestis* dynamics that has yet to be incorporated into these predictive models. The central question is then, what property of these burrows is not anticipated by the present model?

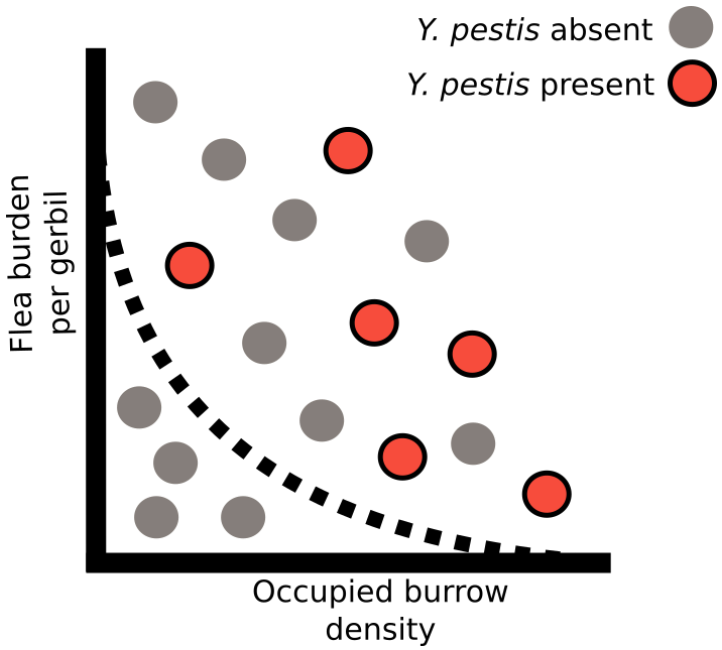


Figure 1.4. Sketch of the output from the curved threshold model, adapted from (Reijniers et al. 2012).

The gerbil population is highly spatially structured, by virtue of their association with semi-permanent burrows. This has been implicit in the previous models of the system, but not explicitly related to empirical records of individual burrows. New data sets collected between 2011 and 2013 detail the individual location of burrows, allowing for closer resolution questions to be considered. Previous work on the data at this scale has identified that not only can burrows be identified from satellite images of the area, but moreover that occupied burrows can be discerned from this (Wilschut et al. 2013). Furthermore, particular geographical structures in the environment can restrict the spread of *Y. pestis* infection across the area, due to restriction of gerbil movement and interaction (Wilschut et al. 2013).

The work presented here continues in this consideration of fine scale interactions and their contribution to the population level dynamics. In the first instance this is considered in terms of the properties of burrows, and whether this can be used to anticipate occupancy patterns. In Chapter two, individual burrow locations and the ecological properties of the surrounding area such as substrate composition and vegetation presence are assessed as predictors for burrow presence and occupancy. It is identified that some environmental variables are significant predictors of the probability of a burrow being, and further of remaining occupied. The text is

presented as in the published manuscript, (Levick et al. 2015). In Chapter three, a metapopulation model of the gerbil population is constructed and the potential for the findings of Chapter two to inform this model explored. This work identifies that neither the metapopulation model, or simply connectivity between occupied burrows, is sufficient to explain occupancy patterns. This was investigated further through the use of simulations to explore the power of the employed analysis, showing that a much larger field data set may be required to reach a satisfactory conclusion.

Chapter four again considers relationships at the individual burrow scale, but in this case of the phenotype of the *Y. pestis* isolates present. Atypical isolates lacking the fraction 1 surface antigen have been identified, and hypothesised to have a role in determining epizootic dynamics. Furthermore, given diagnosis of *Y. pestis* infection in *R. opimus* is often conducted via testing for seroconversion status to the F1 antigen, this may be affecting our ability to estimate prevalence. Were a systematic relationship to be identified, this could be used to correct existing prevalence estimates. Records detailing *Y. pestis* phenotype from across the study region are used to show the atypical strains are commonplace and widespread geographically. Whilst only appearing in burrows following previous typical strains, no clear relationship could be identified between their presence and epizootic dynamics.

1.2 Onchocerciasis associated epilepsy in the Democratic Republic of Congo.

Onchocerciasis is a parasitic disease of humans, caused by infection with the filarial roundworm *Onchocerca volvulus*. The worms are transmitted by black flies (*Simulium* spp.), generally *Simulium damnosum s.l.* in Sub Saharan Africa. Microfilariae are injected into the bloodstream as the black fly feeds on a human host, from where they migrate to the subcutaneous tissue where they form nodules within which they mature. On maturation the dieoceous adults generate microfilariae through sexual reproduction. The released microfilariae can establish further nodules in the host, allowing for more reproduction, or may be acquired by a subsequent feeding black fly (Figure 1.5). Adults can survive in the host for an average of 10 years, occasionally reaching twenty or thirty (Simonsen 2009). Onchocerciasis is often referred to as “river blindness”, with one its starkest symptoms being partial or complete loss of vision due to the accumulation of scar tissue on the cornea or sclera of the eye, arising from immune attacks against microfilariae migrating to the tissues.

Blindness is one of several possible symptoms, generally representing the severe manifestation. Similar immune responses to worms in the subcutaneous tissue cause a range of skin damage. Further, the infection can cause a range of systemic responses, causing weakness and loss of body mass. More recently it has been identified that a bacterial endosymbiont of *O.volvulus*, *Wolbachia pipientis* not only contributes to the fitness of the worm, but also to the pathogenic reactions observed (Brattig 2004). Perhaps paradoxically to the damage caused by hyperimmune reactions, the parasite is also known to immunomodulate the host (Schönemeyer et al. 2001), leaving them vulnerable to a range of further infectious diseases.

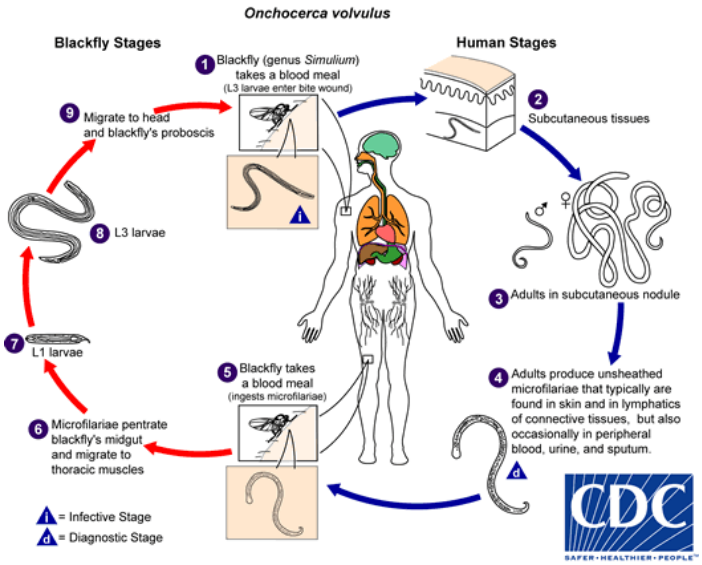


Figure 1.5. Life cycle of *Onchocerca volvulus* (CDC 2013).

Two large multi country control programmes were established to reduce the prevalence of onchocerciasis in Africa; the Onchocerciasis Control Programme (OCP) and the African Programme for Onchocerciasis Control (APOC). The OCP was established in the 1970's, focussing initially on larviciding. This formed a major limiting factor in its roll out to other countries, as it was not possible to conduct this control in densely forest areas. Following the identification of the drug Ivermectin, and its action against a number of roundworms including *Onchocerca spp.* (Campbell et al. 1983), community treatment studies demonstrated the drug was effective in clearing human infection, and moreover interrupting transmission, despite rare events of adverse reactions (de Sole et al. 1989). Following a donation programme established by Merck of their formulation Mectizan, Ivermectin administration was included as an addition to these larviciding campaigns. The APOC was established

later in 1995 in a further 19 countries. APOC relies heavily on the establishment of Community Directed Treatment with Ivermectin (CDTI) initiatives in country, with local staff selected by communities administering Ivermectin doses to eligible individuals (individuals aged >5 years, with no concurrent loaisis infection and non-pregnant women). The OCP in particular is heralded as a large success within public health, believed to have prevented over 600,000 cases of blindness and allowed farming to return to 25 million hectares of land (Boatin 2008). Unfortunately this is not universal, and it has been suggested that these successes may even have held back potential progress elsewhere, perhaps through donor or public interest fatigue (Basáñez et al. 2006). CDTI in particular relies upon spatially and temporally consistent administration as whilst effective individually for around twelve months (Goa et al. 1991), in order to reduce transmission high coverage is required for a minimum of twelve years. Furthermore, it has been suggested latterly that this may be an underestimate, and that realistic targets (less than twenty five years continuous administration) could only be achieved through biannual administration of the drug (Katabarwa & Richards 2014). This has been hard to maintain in many countries, especially those experiencing unrest and conflict. As such, whilst there are many successes, several countries still have large numbers of individual affected by onchocerciasis.

In contrast, epilepsy is a non-communicable chronic neurological disorder, defined by an individual experiencing a minimum of two unprovoked seizures. It can be caused by a number of aetiologies such as congenital defects or damage to the brain arising through traumatic experiences such as strokes or tumours, or the presence of a number of infective organisms. Around 50 million people worldwide are affected by epilepsy, but 80% of them live in low and middle income countries. Treatment rates are low in these countries: only 10% of affected individuals in Africa are thought to receive treatment. Worldwide, people with epilepsy still face a great deal of stigma and misunderstanding (WHO 2015). In Sub Saharan Africa, a great deal or morbidity and mortality in people with epilepsy arises through accidents, such as falling into fires and rivers during seizures (WHO 2002).

A correlation between epilepsy and onchocerciasis prevalence has been reported from a number of studies across Africa, although it is often not definitive and remains without a known functional relationship (Pion et al. 2009). Alongside this, an

atypical form of epilepsy is observed exclusively in these areas. Referred to as "nodding syndrome" (NS), affected patients experience atonic seizures characterised by a dropping motion of the head (Dowell et al. 2013). The exact aetiology of nodding syndrome, like the concurrence of epilepsy with onchocerciasis, remains uncharacterised. Nodding syndrome in particular has been strongly associated with onchocerciasis, in matched studies in Uganda and South Sudan, but again the picture is not clear (Wamala et al. 2015). Notably, infectious diseases known to cause epilepsy, such as cystercercosis and malaria, do not provide sufficient explanation for all affected individuals (Druet-Cabanac et al. 1999). Demographically, individuals with NS do not appear to differ significantly from individuals with other convulsive disorders. Anti convulsive drugs have been shown to be effective at reducing seizure frequency in individuals with NS, accompanied by improvements in other indicators such as a reduction in behavioural and emotional problems, and increased self care (Idro et al. 2014).

The present body of work utilises a data set detailing 12,408 individuals from households across the Orientale province to the North East of the DRC. The data was collected with the primary aims of obtaining good estimates of prevalence of epilepsy, and identify putative risk factors and drivers for epilepsy and nodding syndrome. Additionally, individuals were recruited to a case control study from the Tshopo and Ituri regions.

Chapter five details this work so far, and their implications on potential causes and drivers of epilepsy in the described populations. In particular, the results from a case control study are presented, demonstrating differences in the medical history and behaviours of individuals with epilepsy. These suggest Ivermectin use and exposure to black flies as being associated with epilepsy risk. This follows into a second, larger prevalence study. The aggregated results of this study contrast that of the case control study, finding no clear link with Ivermectin use and epilepsy. However analysis of the data using spatial techniques suggests that this may be explained by patterns of local behaviour, that will require much further investigation to definitively characterise.

Chapter six discusses the key findings from the plague and epilepsy studies, and considers the role of ecological drivers in the understanding of diseases of medical and veterinary health relevance. Ultimately, the following studies highlight

the manner in which infectious disease dynamics are intimately connected to environment in which hosts live, and that public health and ecology form two strands of potential information for controlling and informing decision making.

1.3 References

Basáñez, M.G. et al., 2006. River blindness: A success story under threat? *PLoS Medicine*, 3(9), pp.1454–1460.

Begon, M. et al., 2006. Epizootiologic parameters for plague in Kazakhstan. *Emerging infectious diseases*, 12(2), pp.268–73.

Bertherat, E., 2009. Weekly epidemiological record: relevé épidémiologique hebdomadaire. *The World Health Organization Weekly Epidemiological Record*, (47), pp.445–452.

Boatin, B., 2008. The Onchocerciasis Control Programme in West Africa (OCP). *Annals of tropical medicine and parasitology*, 102 Suppl(1), pp.13–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18718148.

Bolker, B. & Grenfell, B., 1995. Space, persistence and dynamics of measles epidemics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 348(1325), pp.309–20.

Brattig, N.W., 2004. Pathogenesis and host responses in human onchocerciasis: Impact of Onchocerca filariae and Wolbachia endobacteria. *Microbes and Infection*, 6(1), pp.113–128.

Campbell, W.C. et al., 1983. Ivermectin: a potent new antiparasitic agent. *Science (New York, N.Y.)*, 221(4613), pp.823–828.

CDC, 2013. https://www.cdc.gov/dpdx/onchocerciasis/. *CDC DPDx*.

Davis, S. et al., 2007. Plague metapopulation dynamics in a natural reservoir: the burrow system as the unit of study. *Epidemiology and infection*, 135(5), pp.740–8.

Davis, S. et al., 2004. Predictive thresholds for plague in Kazakhstan. *Science (New York, N.Y.)*, 304(5671), pp.736–8.

Davis, S. et al., 2008. The abundance threshold for plague as a critical percolation phenomenon. *Nature*, 454(7204), pp.634–7.

Derbise, A. et al., 2015. Complete Protection against Pneumonic and Bubonic Plague after a Single Oral Vaccination. *PLOS Neglected Tropical Diseases*, 9(10), p.e0004162.

Dowell, S.F. et al., 2013. Nodding syndrome. *Emerging Infectious Diseases*, 19(9), pp.1374–1384.

Druet-Cabanac, M. et al., 1999. Onchocerciasis and epilepsy: a matched case-control study in the Central African Republic. *American journal of epidemiology*, 149(6), pp.565–570.

Eads, D.A. et al., 2016. Droughts may increase susceptibility of prairie dogs to fleas: Incongruity with hypothesized mechanisms of plague cycles in rodents. *Journal of Mammalogy*, 97(4), pp.1044–1053.

Eisen, R.J. & Gage, K.L., 2009. Adaptive strategies of yersinia pestis to persist during inter-epizootic and epizootic periods. *Veterinary Research*, 40, pp.1–14.

Gage, K.L. & Kosoy, M.Y., 2005. Natural history of plague: Perspectives from More than a Century of Research. *Annual Review of Entomology*, 50(1), pp.505–528.

Goa, K.L., McTavish, D. & Clissold, S.P., 1991. Ivermectin: A Review of its Antifilarial Activity, Pharmacokinetic Properties and Clinical Efficacy in Onchocerciasis. *Drugs*, 42(4), pp.640–658.

Idro, R. et al., 2014. Patients with nodding syndrome in Uganda improve with symptomatic treatment: a cross-sectional study. *BMJ open*, 4(11), p.e006476.

Katabarwa, M. & Richards, F., 2014. Twice-yearly ivermectin for onchocerciasis: The time is now. *The Lancet Infectious Diseases*, 14(5), pp.373–374. Available at: http://dx.doi.org/10.1016/S1473-3099(14)70732-7.

Kausrud, K.L. et al., 2007. Climatically driven synchrony of gerbil populations allows large-scale plague outbreaks. *Proceedings of the Royal Society: Biological Sciences*, 274(1621), pp.1963–9.

Kausrud, K.L. et al., 2010. Modeling the epidemiological history of plague in Central Asia: palaeoclimatic forcing on a disease system over the past millennium. *BMC biology*, 8, p.112.

Levick, B. et al., 2015. The perfect burrow, but for what? Identifying local habitat conditions promoting the presence of the host and vector species in the Kazakh plague system. *PLoS ONE*, 10(9), p.e0136962.

Lloyd-Smith, J.O. et al., 2005. Should we expect population thresholds for wildlife disease? *Trends in Ecology and Evolution*, 20(9), pp.511–519.

Naumov, N.P. & Lobachev, V.S., 1975. Ecology of desert rodents of the USSR (Jerboas & Gerbils). In I. Prakash & P. K. Ghosh, Eds. . *Rodents in Desert Environments*. The Hague: The Hague, pp. 465–598.

Perry, R.D. & Fetherston, J.D., 1997. Yersinia pestis--etiologic agent of plague. *Clinical microbiology reviews*, 10(1), pp.35–66.

Pion, S.D.S. et al., 2009. Epilepsy in Onchocerciasis Endemic Areas: Systematic Review and Meta-analysis of Population-Based Surveys. *PLOS Neglected Tropical Diseases*, 3(6), p.e461.

Randall, J.A. et al., 2005. Flexible social structure of a desert rodent, Rhombomys opimus: philopatry, kinship, and ecological constraints. *Behavioral Ecology*, 16(6), pp.961–973.

Reijniers, J. et al., 2012. A curve of thresholds governs plague epizootics in Central Asia. *Ecology letters*, 15(6), pp.554–60.

Riehm, J.M. et al., 2015. Diverse Genotypes of Yersinia pestis Caused Plague in Madagascar in 2007. *PLOS Neglected Tropical Diseases*, 9(6), p.e0003844.

Schmid, B. V. et al., 2015. The pulse of Asia: the Black Death and successive reintroductions of plague into Europe. *Proceedings of the National Academy of Sciences of the United States of America*, Early Edit, pp.1–6.

Schönemeyer, A. et al., 2001. Modulation of human T cell responses and macrophage functions by onchocystatin, a secreted protein of the filarial nematode Onchocerca volvulus. *Journal of immunology (Baltimore, Md. : 1950)*, 167(6), pp.3207–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11544307.

Simonsen, P.E., 2009. Filariases. In G. C. Cook & A. I. Zumla, eds. *Manson's Tropical Diseases*. pp. 1491–1500.

de Sole, G. et al., 1989. Adverse reactions after large-scale treatment of onchocerciasis with ivermectin: combined results from eight community trials. *Bulletin of the World Health Organization*, 67(6), pp.707–719.

Stenseth, N.C. et al., 2008. Plague: past, present, and future. *PLoS medicine*, 5(1), p.e3.

Wamala, J.F. et al., 2015. Nodding Syndrome May Be Only the Ears of the Hippo. *PLOS Neglected Tropical Diseases*, 9(8), p.e0003880.

WHO, 2002. EPILEPSY : A manual for Medical and Clinical Officers in Africa and Clinical Officers. *World Health Organisation*, p.124p.

WHO, 2016. Global distribution of natural plague foci as of March 2016. *Plague: technical information.*, (March), p.2016. Available at: http://who.int/csr/disease/plague/Plague-map-2016.pdf?ua=1 [Accessed November 26, 2016].

WHO, 2015. *Kit on Epilepsy What you can do*, Available at: http://www.who.int/mental_health/neurology/epilepsy/information_kit_epilepsy/en/.

Wilschut, L.I. et al., 2013. Mapping the Distribution of the main host for plague in a complex lanscape in Kazakhstan : An object-based approach using SPOT-5 XS , Landsat 7 ETM + , SRTM and multiple Random Forests. *International Journal of Applied Earth Observation and Geoinformation*, 23, pp.81–94.

Wilschut, L.I. et al., 2013. Potential corridors and barriers for plague spread in Central Asia. *International journal of health geographics*, 12(1), p.49.

Zeppelini, C.G., de Almeida, A.M.P. & Cordeiro-Estrela, P., 2016. Zoonoses As Ecological Entities: A Case Review of Plague. *PLOS Neglected Tropical Diseases*, 10(10), p.e0004949.

2. The Perfect Burrow, but for what? Identifying local habitat conditions promoting the presence of host and vector species in the Kazakh plague system.

2.1. Introduction

Despite the relative rarity of plague (*Yersinia pestis* infection) in modern human populations, historical outbreaks such as the Black Death across medieval Europe clearly demonstrate the disease's devastating potential (Stenseth et al. 2008). Now largely confined to foci in wildlife populations, plague has been characterised as an emerging (or re-emerging) infectious disease (Morse 1995). Emerging infectious diseases and in particular zoonotic infections from wildlife species are considered a great future danger to human health worldwide (Daszak 2000), with their emergence being linked to both multiple socio-economic factors and environmental conditions (Jones et al. 2008). Being able to increase our ability to predict and prevent epizootics, and in particular, minimising the risk of spillover into human populations, is essential for controlling these emerging diseases.

The *Y. pestis* focus in the Pre-Balkhash desert of Kazakhstan has been subject to extensive study over the past 15 years. The desert and semi-desert belt across the region consists of a vast area of mixed vegetation and shrubs that extend from the Caspian Sea in the west to the Chinese border in the south east of the country, in the Almaty province. The northern desert sub-zone of the Almaty province is the largest of 20 endemic plague foci in the country. The hosts from this focus have been recently suggested to have had involvement in initiating the European outbreaks of plague in the 1300's (Schmid et al. 2015). However, since the 1950's the national (previously USSR) anti-plague system has been largely successful in preventing the occurrence of human cases (Onishchenko & Kutyrev 2004; Rivkus et al. 1992). Across 14 of the foci located in the Pre-Balkhash region, plague is mainly transmitted between great gerbils (*Rhombomys opimus*) by a variety of fleas, although *Xenopsylla* spp. generally dominate. *R. opimus* is the dominant rodent in the landscapes of the Pre-Balkhash area, inhabiting complex burrow systems as extended family groups. Previous work has identified that plague spreads through the gerbil population in a manner well described by percolation theory (Davis et al. 2007). Importantly, this model followed previous work (Davis et al. 2004) in predicting a critical threshold of gerbil abudance (defined as the proportion of gerbil burrow systems that are occupied) necessary for a plague epizootic to occur. A more recent adaptation of the model

extends this to describe the threshold as a curve arising as a function of both the abundance of the gerbils and their flea burden (Reijniers et al. 2012). However, the dynamics of plague transmission, and of the gerbil population, have not been considered previously at the scale of the individual burrow.

R.opimus burrows are visible on satellite images due to depletion of vegetation in the active (or so- called ecological) centre of the burrow, and the surrounding area where activities are most intense (Davis et al. 2007; Addink et al. 2010;Wilschut et al. 2013). Like many other diurnal rodents, their complex and extensive burrow systems are important for their social behaviours (Randall et al. 2005) and provide nesting and food chambers, safe access to above ground food resources and refuges from predators (Shenbrot et al. 2002). The great gerbil's burrowing activities promote the "fertile island effect" in influencing the concentration of soil nutrients and available nitrogen (Xu et al. 2012) and as such the great gerbil can be considered as a powerful ecosystem engineer, capable of modifying physio-chemical properties of the soil at the fine burrow system scale. These burrows also provide a buffered environment and refuge for a number of invertebrate species, in particular those that feed on gerbils and are potential vectors (mainly fleas and ticks) of pathogenic agents. The number, position, and size of the burrow systems generally do not change over time, but the proportion of burrow systems occupied by family groups may fluctuate dramatically, hence making burrow occupancy an effective proxy for great gerbil abundance (Davis et al. 2007; Davis et al. 2004).

The aim of this investigation is to improve the ability to predict plague incidence in the Pre- Balkhash region, through increased ability to predict the presence of the rodent host and the abundance of the flea vector at the level of the individual burrow. Given their importance for the presence of plague, we can consider that a burrow that is highly preferred by both the host and vector may be a perfect burrow for *Y.pestis*. Further, for plague to spread across a landscape, burrows occupied by the host must be distributed with sufficient connectivity to permit percolation through the population. Preference for burrows with specific attributes may cause clustering of burrows supporting gerbils and fleas within the landscape, with extensive gaps in between, potentially preventing long distance spread of plague (Kausrud et al. 2007).

Descriptive accounts of *R.opimus* burrow properties and ecological preferences have been reported previously (Naumov & Lobachev 1975). More recent work has identified landscape-scale properties with functional corridors and barriers for potential plague transmission; that is, areas of high and low density of burrows (though all burrows were considered rather than those occupied by gerbils or fleas; (Liesbeth I Wilschut et al. 2013)). However, as yet there has been no quantitative analysis of landscape features at the level of the individual burrow that are associated with their being occupied, with remaining occupied over a number of seasons, or with gerbils in the burrow having a high flea burden. Such is the aim of the present study.

2.2. Aims and hypotheses

Aim

- Investigate whether environmental properties of individual *R. opimus* (great gerbil) burrows in Kazakhstan are predictive of the presence of *R. opimus* or *Xenopsylla* fleas.

Hypotheses

- Environmental properties of a burrow are related to the presence of *R. opimus*.
- Environmental properties of a burrow are related to the presence of *Xenopsylla* fleas.

2.3. Methods

Study site

The study sites are located in the Pre-Balkhash endemic plague focus in south eastern Kazakhstan (ranging between 44°75′43.61 and 45°73′73.80; see also (Wilschut et al. 2013)). The cartographic division of the territory for plague monitoring into primary squares (40km x 40km), and then each into four secondary squares (20km x 20km) and these into four sectors (10km x 10km), as well as the large scale features of the landscape, have been described elsewhere (Davis et al. 2007; Addink et al. 2010; Wilschut et al. 2013; Stenseth et al. 2006). The study sites were located in 6 sectors on the 88 east bank of the Ili River within the alluvial plains that follow the historical delta in the northern portion 89 of the focus (Gvozdetskii & Mikhailov 1963). The semi-desert vegetation consists of short grass (Gramineae, Poaceae), herbs (mainly 90 Chenopodiaceae, halophytic vegetation) and shrubs dominated by *Haloxylon* and *Calligonum* species.

In four seasons (monitoring periods of around 1-2 months; autumn (September) 2011, summer (July) 2012, autumn (September) 2012 and spring (April-May) 2013), all burrow systems located within randomly chosen 500m x 500m (25ha) squares spread over the 6 sectors were mapped. In total, 2113 burrows were mapped. Of these, occupancy status was recorded twice (in two seasons) for 894 burrows and just once for 1219 burrows. The GPS location of the burrow system was recorded along with its altitude and its occupancy status.

Individual burrow systems were identified by locating the central burrow structures and the outer limits of each discrete burrow unit (typically around 50m in diameter). For each burrow a description of its occupancy status and the landscape of the area surrounding the burrow was created. If direct evidence of gerbils was found at the burrow (fresh faeces, obvious digging) then it was classified as occupied. Signs of gerbil activity limited to feet/tail markings in the sand at the entrances of burrows would classify a burrow as visited. If no gerbil activity was identified then the burrow was classified by its condition; if in usable condition but otherwise not inhabited the burrow was classified as unoccupied, varying states of disrepair would identify the burrow as old or disappeared.

A landscape description of the burrow was then generated. This was based on visual inspection of landscape properties of the local area in which the burrow had been constructed, in particular guided by granulometry to classify sediments. Classification of landscape properties was thus conducted by methods detailed in Pogrebinsky (1963) (Gvozdetskii & Mikhailov 1963; Pogrebinsky 1963) and those specifically relating to the granulometry techniques are detailed further in Wentworth 1992 (Wentworth 1992). The properties, recorded for each burrow, were split into sediment properties of which there were eight (presence of sand, clay etc.), and topographical features of which there were seven (position on a dune, waterlogged areas, etc.). These characteristics and their recording is detailed elsewhere (Levick et al. 2015: File S1).

During daylight hours, great gerbils (*R.opimus*) were caught using local traps placed at burrow entrances; trapping procedures and laboratory protocols were performed according to Russian standard methods (Burdelov 1998; Korneyev 2005). Gerbil traps were visited frequently so as to minimise their time in a trap (typically

less than thirty minutes). Following this they were euthanised according to Kazakh public health protocols, as approved by the University of Liverpool Animal Welfare Committee. Trapped rodents were brought to the laboratory for processing. Individual rodents were sexed, weighed, and classified as juvenile or adult. For each rodent, fleas and ticks brushed off rodent fur were identified to species (fleas) or genus (tick) level, counted, sexed, and crushed in saline (NaCl 1%), before plating on selective agar to attempt to isolate *Y.pestis*. Further details of the laboratory techniques can be found elsewhere (Levick et al. 2015: File S1). The number of such fleas collected from a rodent trapped in a burrow divided by the number of rodents trapped from the burrow is referred to as the flea index (and tick index was computed similarly).

A second data set (hereafter referred to as the long term data set) was collected to examine occupancy over multiple seasons. Methods differed little from those above except that the squares monitored were 200m x 200m, and rodents were not trapped. 513 unique burrows were revisited in each of six seasons (spring, summer and autumn 2011; spring and summer 2012; spring 2013) and their occupancy recorded each time. Some of these burrows were trapped as per the protocol above in Spring 2013 and these 129 records added to the main data set.

Generating landscape factors

A unique burrow landscape description was generated for each of the 2113 mapped burrows. Some of the landscape characteristics are mutually exclusive, but many can occur together and so as many characteristics as were needed to fully describe the burrow were noted in each case. The resulting data set was comprehensive and highly detailed, but the large variety (171 unique descriptions) of individual burrow landscape descriptions required condensing in order to identify key aspects of the landscape linked with the outcome variables (occupancy, long term maintenance of occupancy, burden of fleas).

Landscape descriptions could be considered at their finest scale (e.g. the burrow is on the side of a dune, on the top, or on the very bottom) or at a coarser level (the burrow is on a dune). There was little *a priori* indication of the coarseness of landscape description relevant to the outcome variable. Therefore, an initial exploratory quantitative analysis, designed specifically for this data set, was conducted to identify general trends in the outcome variables across different

landscape descriptions and groups of characteristics. For the two general outcome variables (occupancy levels, fleas) the level of an outcome response associated with each of the burrow landscape descriptions across the data set was estimated. For fleas this was measured as the mean flea index associated with burrows possessing each landscape description, and for occupancy as the proportion of burrows possessing this description that were occupied. In each case this was normalised by the number of burrows possessing each description to avoid over-representation of the effect of burrow properties that were simply common throughout the landscape.

All of the unique individual burrow descriptions in the data set were ranked by their level of outcome response (normalised by the number of burrows possessing this description). The ranked descriptions were then split into a high, medium and low group by their associated response level. These groups each contained roughly one third of the descriptions, with some allowance for groups of descriptions with the same level of response to be grouped together. For each of these groups, the proportion of descriptions in the group containing each of the 15 characteristics was calculated. Where the proportion of descriptions containing a particular characteristic differed by approximately 5-20% between the high and medium or low and medium groups, the characteristic was considered either as a predictive landscape factors to be used in statistical analysis or for inclusion within a predictive factor (Levick et al. 2015: File S1). In the latter case, where landscape characteristics describing related properties showed similar relationships with the outcome variable, they were grouped together to form landscape factors to be used in the statistical analysis. One such case, for example, were the landscape descriptors relating to the burrow being located on a dune (D,E,F,7 & 8), including descriptions of the position on the dune and the stability of the dune structure, but which were ultimately grouped together into the ”dune” factor. The resulting landscape factors, overall, are detailed in Table 2.1.

Table 2.1. Landscape factors used in models, built using landscape descriptors listed above.

Model variable	Nature of variable
Sand Clay Loam Solonchak/takir Dunes	Binary, presence or absence of indicated property

Statistical Models

Generalised linear mixed models (GLMMs) were used to explore the relationship of the model variables described in Table 2.1 with both occupancy status and flea burden. In general, the models were used to test hypotheses regarding putative predictors identified from field observations and the biology of the system, detailed below in Table 2.2.

Table 2.2. Landscape factors and further environmental factors tested in each model.

Response variable	**Landscape factors tested**	**Environmental factors tested**
Occupancy	Solonchak, sand, loam, clay, dunes/clay ∗ dunes	Latitudinal position, season, tree presence
Long term occupancy	Solonchak, sand, loam, clay, dunes/clay ∗ dunes	Latitudinal position, season, tree presence
Flea burden	Sand, loam, clay, dunes	Latitudinal position, season, tree presence, tick index, burrow occupancy in surrounding sector that season
Xenopsylla burden	Sand, loam, clay, dunes	Latitudinal position, season, tree presence, tick index, burrow occupancy in surrounding sector that season

In particular for models of occupancy, apart from individual landscape factors, the interaction (and respective main effect terms) between burrows being on clay sediment and the presence of dunes was included to test the hypothesis that their potential advantages (for clay sediment: stability; for dunes: ease of digging, an angled entrance providing extra shelter from wind/rain) would make their combined presence of greater benefit. Tree presence was included to test the hypothesis that trees are indicators of moisture concentration in the environment, and therefore in turn indicate fertile areas with ready food supply. For models of flea burden, the tick index was tested as a predictor to test whether certain burrows were more likely to be generally infested with invertebrates. In addition the percentage of burrows occupied

in the surrounding sector (in the same season as the flea record was taken) was tested as a predictor of flea index to identify whether the flea burden was related to gerbil abundance.

Predicting Occupancy

The field recording of occupancy status was condensed to a binary record of occupied or unoccupied (the latter including all other recorded statuses; visited, empty or old). GLMMs using a binomial error distribution were constructed to predict occupancy for the full data set of 3011 observations (i.e. trapping events) of 2113 unique burrows, including repeat visits. Additionally, as multiple visits to sectors were not randomised, both burrow identity and sector were included as random effects (with burrow identity nested by their relevant sector).

A second GLMM was constructed to identify factors predictive of burrows remaining occupied using the long term occupancy data set, using, as the independent variable, the proportion of the six monitoring periods where each of the revisited burrows was occupied. As above, sector was included as a random effect. The same selection of landscape factors was used in both cases.

Predicting flea numbers

The data set contained 999 observations of flea burden, of which *Xenopsylla spp.* dominated as anticipated (making up 98.9% of the observations). A GLMM was constructed using records of flea burden, using a poisson error distribution, and using a quasi-poisson model as a correction of standard errors due to over dispersion in the data (Ver Hoef & Boveng 2007), while including burrow and sector as nested random effects (as in the static occupancy model above).

Model reduction and comparison

In each case, a model was first constructed using all of the burrow level landscape traits hypothesised to be related to the outcome. This model was minimised as detailed below. A second model was then built using the significant landscape factors with the addition of further environmental factors, as detailed in Table 3. Again, this was minimised by the same process. Akaike's Information Criterion (AIC) scores were estimated for the occupancy and long term occupancy models by finding the AIC score for the non-mixed version of the model (i.e. without random effects).

For the models with quasipoisson error corrections, a quasi-AIC (qAIC) score was estimated, using a non mixed version of the model to obtain a dispersion parameter, and a model without error corrections to obtain a likelihood score (Bolker et al. 2009). Models with a ΔAIC of less than two were considered to have equal explanatory ability (Burnham & Anderson 2002). Where all factors were significant, backwards elimination was applied to identify any significant factors whose removal may improve the model AIC score. This was not performed in the flea models due to the additional complexity (and associated uncertainty) in the methods used to obtain the AIC estimates as detailed above.

However, while the AIC scores provide a method of comparing the different models, in each case they are estimated for the non-mixed version of the model and so we consider them to be approximate measures of the fits of the models with random effects included. They were therefore considered alongside the complexity of the model and the significance of the parameters. Specifically, simpler models, and those where all terms were significant, notwithstanding AIC score, were prioritised as detailed below.

In the final parsimonious static occupancy and flea models, checks of the nested random effects structure were performed by iteratively running the model, using a subset of the data consisting of one record from every instance where a burrow was sampled more than once, alongside all records of burrows only recorded on one occasion. The range and average of the coefficients and P values in these models were compared to those produced by the models using the nested random effects (see also (Telfer 2010)). These results are shown in the supplementary material. Only where effects from the original models were confirmed by this alternative were results considered reliable (details below).

All statistical analysis was conducted using the R statistical computing environment (R Core Team 2014), using the additional package MASS (Venables & Ripley 2002).

2.4. Results

Occupancy status

Of the terms included in the landscape factor model, the interaction term between clay sediments and dunes was found to be significant. However the ΔAIC between the models with and without the interaction term did not differ significantly (AIC with interaction=4032.86, without interaction=4032.74). Hence, the simpler model without the interaction term was chosen as a minimal landscape model. Within this only the clay sediment term was significant and this was carried forward to the second model. Of the additional terms then tested, latitudinal position, the presence of trees and season were found to be significant. Backwards selection did not suggest removal of any of these would create significant improvement in the AIC scores, and as all terms were significant this was maintained as the minimal model. All of the model terms were validated by the iterative sampling method (Levick et al. 2015: File S1). Descriptions of these relationships, and their associated coefficients are detailed in Table 2.3.

Table 2.3. Description of terms in the parsimonious model of occupancy status.

Factor	Description of relationship	Coefficient	P Value	Standard Error
Latitudinal position	Occupancy higher with increasing latitude (to the north)	2.14	<0.01	0.41
Tree presence	Occupancy higher in the presence of trees	1.19	<0.01	0.12
Season (Summer 2012)	Occupancy lower than the reference of Autumn 2011	-0.35	<0.03	0.16
Season (Autumn 2012)	Occupancy lower than the reference of Autumn 2011	-0.45	<0.05	0.18
Season (Spring 2013)	Occupancy lower than the reference of Autumn 2011	-0.95	<0.01	0.11
Clay	Occupancy lower in presence of clay	-0.42	<0.01	0.13

Long term occupancy

The same landscape factors were tested as in the occupancy model. The interaction term between clay sediment and dunes, and its constituent main effects were found to be significant alongside sand. Here the model with the interaction had a significantly lower AIC (without interaction=2155.80, with interaction=2090.71) compared to the simpler model with just the main effects, and so as the terms were significant this was taken as the minimal landscape model. Of the additional environmental factors tested alongside these landscape factors, only latitude was found to be significant. The model including the non-significant environmental terms (season, presence of trees) was significantly better in terms of AIC (1625.00 with insignificant terms, 1770.00 without), but the reduced model was selected due to its increased simplicity and exclusion of insignificant terms. Again, backwards selection did not indicate that removal of any of these terms would significantly improve the model. Descriptions of these relationships, and their associated coefficients are detailed in Table 2.4.

Table 2.4. Description of terms in the parsimonious model of multiple season occupancy status.

Factor	Description of relationship	Coefficient	P Value	Standard Error
Latitudinal position	Burrows in northern regions occupied for a higher proportion of seasons	3.55	<0.01	0.44
Sand	Burrows occupied for a higher proportion of seasons in the presence of sand	0.67	<0.01	0.68
Dunes	Burrows occupied for a higher proportion of recorded seasons in the presence of dunes	0.23	0.25	0.2
Clay	Burrows occupied for a higher proportion of recorded seasons in the presence of Clay	0.41	0.12	0.26
Dunes ∗ Clay	Burrows occupied for a lower proportion of recorded seasons when both are together	-1.73	<0.01	0.49

Flea burden

Of the four landscape factors tested, only clay was found to be a significant predictor. Of the additional environmental factors alongside clay, only the tick burden and season were found to be significant. The model including all of the additional environmental factors (latitudinal position, tree presence, sur rounding burrow occupancy) had a significantly lower AIC (with all factors=568, with only significant factors=571), but the simpler model with only significant parameters was taken as the final parsimonious model as detailed in Table 2.5. Again, all of the model terms were validated by the iterative sampling method (Levick et al. 2015: File S1).

Table 2.5. Description of terms in parsimonious model of *Xenopsylla* burden.

Factor	Description of relationship	Coefficient	P Value	Standard Deviation
Tick burden	Xenopsylla burden higher where tick burden is higher	0.02	<0.01	0
Season (Summer 2012)	Xenopsylla burden higher than in reference season of Autumn 2011	0.16	0.12	0.01
Season (Autumn 2012)	Xenopsylla burden lower than in reference season of Autumn 2011	-0.31	<0.01	0.01
Season (Spring 2013)	Xenopsylla burden lower than in reference season of Autumn 2011	-0.02	0.78	0.09

2.5. Discussion

The results presented above demonstrate that both static gerbil occupancy, and the maintenance of burrow occupancy in the long term are related to the burrow's latitudinal position. Further, static occupancy is found to be related to the presence of trees, clay soils and the season, and the maintenance of occupancy is related to the presence of sand, and an interaction between the presence of clay sediments and dunes.

Previous research examining burrow location preferences and burrow construction in a range of fossorial mammals has identified connections between sediment properties and both the resulting burrow structure and the likelihood that a burrow is occupied. In particular, at the genus level members of the Geomyidae family harbour preferences for burrows with particular environmental properties; for example, while *G. busarius* prefer sandier loam sediments, *T.bottae* and *P.castanops* prefer clay-loam sediments (Best 1973; Roberts et al. 2014; Wilkins & Roberts 2007). Additionally, descriptive analysis of other species' burrows show distinct changes in burrow architecture with changing composition of the sediment in which it is constructed. Individuals of the *Dipodomys* genus (kangaroo rats) construct longer, more complex burrows where there are higher percentages of silt, and *Microtus*

montanus (Montane vole) construct deeper, narrower burrows in firmer sediments with more sand (Laundré & Reynolds 1987). Other studies have described cases where previously suitable habitat has been altered for agriculture (thus changing the sediment properties) and has subsequently been avoided by fossorial mammals (Hoffman et al. 2007).

The results above demonstrated that when observed in a single season, burrows in clay sediments are less likely to be occupied. Digging new burrows incurs a cost in any fossorial species, but over the lifespan of the animal this is generally low in comparison to the resultant benefits created by the burrow (Reichman & Smith 1990). The option of constructing a new burrow in clay sediments may be less preferable due to the initially high levels of energy expenditure, despite the stability and high level of protection that one would expect to be conferred by the clay sediments.

In contrast, although neither clay nor dunes on their own had a significant relationship with long term occupancy, the significant interaction term between them indicated that burrows in clay sediments are less likely to remain occupied for longer periods when there are also dunes present, and conversely that burrows in areas with dunes are less likely to remain occupied where clay is also present. The potential benefits one might imagine of sturdier clay soils, and those observed in previous studies (Laundré & Reynolds 1987) may become discounted through disruption in the surrounding landscape created by dunes structures. Similarly, the anticipated benefits conferred by dunes structures (ease of digging due to sand content, and sheltering due to an angled entrance) are both not significantly related to long term occupancy, and where clay sediment is also present any positive effect is mitigated. As dunes do occur frequently across the clay sediment surfaces, the observation that burrows constructed where both are present are less likely to remain occupied has implications for the long term spatial structure of the gerbil populations.

A significant positive association between a burrow being located in an area of sand and long term occupancy was also observed. In *Meriones crassus* burrows constructed in sediments with artificially controlled proportions of sediment types, those constructed in the presence of high proportions of sand result in a higher average temperature of the burrow microclimate compared to other substrate

compositions, suggesting sand plays a role in the insulation of the burrow (Shenbrot et al. 2002). These results then may indicate that this insulating property may be key to over-winter or at least multi-season survival for the gerbils, in the face of the drastic temperature fluctuations (hot summers and cold winters) observed in the Pre-Balkhash. The significant relationship with latitude may reflect the effect of an unmeasured environmental variable that changes along a latitudinal gradient (such as climate), or indeed unmeasured isotropic spatial variation. Gerbil occupancy across the region is known to fluctuate (Davis et al. 2004), but the spatial correlation drops from around 0.7 at distances of around 20km to around 0.55 at 250km (Kausrud et al. 2007). Further the climate in the pre-Balkhash region correlates closes with latitudinal gradient (Deom et al. 2012). Therefore both latitudinally varying factors and latitudinally-independent asynchronies may both play a role here.

The positive relationship between the presence of trees and static occupancy is perhaps suggestive of a necessary level of moisture (supporting both tree growth and occupancy) and/or protection from excess heat through shading, the latter fitting with previous observational records (Naumov & Lobachev 1975). In addition, an experimental study of the vole, *Microtus pinetorum*, identified preference for moderate soil moisture, likely to be a balance between digging ease and risk of flooding (Rhodes et al. 1985). Albeit in different ecosystems, small mammals have been shown to positively impact the dispersal of tree seeds throughout the environment (Brewer & Rejmánek 1999). Bioturbation of the soil by gerbils as they burrow may also aid plant growth by enabling redistribution of micronutrients in the soil (Whitford & Kay 1999). Thus, the gerbils themselves may increase the presence of trees, possibly creating a positive feedback system or so called “fertile islands” (Xu et al. 2012).

Lower occupancies in spring in particular are likely simply to reflect the natural fluctuations of gerbil populations, where numbers are generally highest in autumn (e.g., (Davis et al. 2004)), following which reproduction ceases and numbers decline over the winter. Furthermore, the lower occupancies in the later trapping seasons may also simply reflect the general decline in gerbil numbers across the entire region during this period. Field observations suggest a gerbil population crash occurred around 2011, but that gerbil numbers remained higher in the Northern region (also likely to influence the relationship with latitude).

Fleas

Although several flea species in this system are potential *Y. pestis* vectors, *Xenopsylla spp.* are the most important (Rapoport et al. 2010), (with *Coptopsylla lamellifer* playing a role in winter transmission) and are the dominant species in our data set. The different species themselves differ in their seasonal dynamics and details of their habitat requirements. Here, latitude, season and tick burden associated with a burrow were identified to be significantly related to the burden of fleas in the burrow.

The effect of season is likely to be linked with the associated abiotic environmental changes. There may also be an impact in the changing vegetation through the seasons, with the vegetation cover determining shade, temperature and soil moisture at scales affecting the fleas (Krasnov et al. 2001). Here, rodents trapped in Autumn 2012 had a lower flea burden. This likely reflects the dip in flea numbers over the colder months typically observed in the *Xenopsylla* life cycle, combined with a general decline in host abundance. In general these relationships reflect the well documented effects of season on *Xenopsylla*. Similarly to many insect species, external temperatures have direct effects on the metabolic and therefore developmental processes of fleas, and their survival is reduced at low humidities (Krasnov et al. 2001; Cavanaugh 1971). For fleas in particular, as developmental stages occur off host, their developmental rate and population dynamics are controlled greatly by abiotic factors (Gage et al. 2008).

That tick burden is a significant predictor of flea burden suggests that there may be burrow conditions not monitored here that create conditions permissive of a wide variety of ectoparasites and that are conducive to the spread of plague. Alternatively, since the flea and tick measurements relate to numbers found on individual gerbils, the association may reflect characteristics of these individuals rather than the burrows in which they were trapped. Certainly, flea infestation is known to affect the metabolism of the mammalian host (Khokhlova et al. 2002). Further, alterations to the host behaviour and immune response by an initial ectoparasite infestation may then make the host more susceptible to multiple ectoparasite infestation (Wikel 2011), even before considering the immunological effects of infection by *Y.pestis* to either the gerbils or fleas. Discerning cause and effect here is difficult, but there is certainly an indication that the wider ectoparasite

and invertebrate communities should be considered in order to properly characterise the dynamics of this plague system. Indeed, several authors report transmission or isolation of *Y.pestis* in *Ixodidae* (*Rhipicephalus sp.* and *Hyalomma sp.*) and *Argasidaee* ticks (in particular, *Ornithodoros tartakovsky*) in Central Asian desert foci (Fedorov 1960; Bykov et al. 1966; Semachko 1981), and this was also the case in the present study.

Finally, migrating fleas may provide insight into the specific role of fleas in determining the timing and size of epizootics. A theoretical study based on the American prarie dog plague system identified a potential role for migrating fleas in enzootic persistence between epizootic events (Buhnerkempe et al. 2011). However, data collected here on fleas aspirated directly from burrows was not considered to be sufficiently consistent for use in a statistical analysis equivalent to that for fleas collected directly from gerbils.

A perfect burrow for gerbils, fleas and Yersinia pestis

At the most basic level, for *Y. pestis* epizootic events to spread and be sustained, both the mammalian and insect host must be present across the landscape. This would then suggest that a burrow in the northern region, with surrounding vegetation, and perhaps located in a sandy areas, would be at the highest risk for plague presence given the significant positive relationships between these properties and host and vector presence identified here. Plague prevalence during the study period was not high enough to test this directly. There may also be other factors that need to be present or absent for a burrow to be permissive of *Y.pestis* itself. The question also remains of where the bacterium survives in interepizootic periods (or indeed, through gerbil population crashes), which may itself be controlled by abiotic environmental conditions. Overall, this investigation gives finer resolution to our ability to predict distributions of both of the key host and vector species in this system, so improving our ability to predict plague epizootic occurrence and behaviour. As a starting point to investigate this, a closer analysis could be conducted of environmental factors linked to specific burrow transition events from season to season.

2.6. References

Addink, E. A. et al., 2010. The use of high-resolution remote sensing for plague surveillance in Kazakhstan. *Remote Sensing of Environment*, 114(3), pp.674–681.

Best, T.L., 1973. Ecological Separation of Three Genera of Pocket Gophers (Geomyidae). *Ecological society of America*, 54(6), pp.1311–1319.

Bolker, B.M. et al., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, 24(Table 1), pp.127–135.

Brewer, S.W. & Rejmánek, M., 1999. Small rodents as significant dispersers of tree seeds in a Neotropical forest. *Journal of Vegetation Science*, 10(2), pp.165–174. Available at: http://doi.wiley.com/10.2307/3237138.

Buhnerkempe, M.G. et al., 2011. Transmission shifts underlie variability in population responses to Yersinia pestis infection. *PLoS ONE*, 6(7).

Burdelov, L.A., 1998. *A guide on emergency field and village prophylaxis of plague.*, Almaty: In Russian.

Burnham, K.P. & Anderson, D.R., 2002. *Model Selection and Multimodel Inference* Second., New York: Springer-Verlag.

Bykov, L.T., Popova, A.S. & Solkolva, A.A., 1966. On participation of Ixodoidea ticks in Mufum Kum plague epizootics. *Proc. Tezisy Dolk. 1. Akarol Soveshch*, 3(Translation from Russian), pp.47–48.

Cavanaugh, D.C., 1971. Specific effect of temperature upon transmission of the plague bacillus by the oriental rat flea, Xenopsylla cheopis. *American Journal of Tropical Medicine and Hygiene*, 20(2), pp.264–273.

Daszak, P., 2000. Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health. *Science*, 287(5452), pp.443–449.

Davis, S. et al., 2007. Plague metapopulation dynamics in a natural reservoir: the burrow system as the unit of study. *Epidemiology and infection*, 135(5), pp.740–8.

Davis, S. et al., 2004. Predictive thresholds for plague in Kazakhstan. *Science (New York, N.Y.)*, 304(5671), pp.736–8.

Deom, J. et al., 2012. Quaternary Evolution of the Human Habitats in the Ili-Balkhash Region from Paleolithic to Modern Times. In *The International Workshop.* Almaty, pp. 49–58.

Fedorov, V.N., 1960. Plague in camels and its prevention in the USSR. *Bulletin of the World Health Organization*, 23, pp.275–281.

Gage, K.L. et al., 2008. Climate and Vectorborne Diseases. *American Journal of Preventive Medicine*, 35(5), pp.436–450.

Gvozdetskii, N.A. & Mikhailov, N.I., 1963. Asiatic region. In *Physical Geography of the USSR*. p. 566.

Ver Hoef, J.M. & Boveng, P.L., 2007. Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology*, 88(11), pp.2766–2772.

Hoffman, J.D., Choate, J.R. & Channell, R., 2007. Effects of land use and soil texture on distributions of pocket gophers in Kansas. *Southwestern association of naturalists*, 52(2), pp.296–301.

Jones, K.E. et al., 2008. Global trends in emerging infectious diseases. *Nature*, 451(7181), pp.990–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18288193 [Accessed March 3, 2013].

Kausrud, K.L. et al., 2007. Climatically driven synchrony of gerbil populations allows large-scale plague outbreaks. *Proceedings of the Royal Society: Biological Sciences*, 274(1621), pp.1963–9.

Khokhlova, I.S. et al., 2002. Energy cost of ectoparasitism: the flea Xenopsylla ramesis on the desert gerbil Gerbillus dasyurus. *Journal of Zoology*, 258(2002), pp.349–354. Available at: http://doi.wiley.com/10.1017/S0952836902001498.

Korneyev, G.A. ed., 2005. *Capture, count and forecast of the number of small mammals and birds in the natural foci of diseases: Guidelines.*, Moscow: In Russian.

Krasnov, B.R. et al., 2001. Development rates of two Xenopsylla flea species in relation to air temperature and humidity. *Medical and veterinary entomology*, 15(3), pp.249–58.

Levick, B. et al., 2015. The perfect burrow, but for what? Identifying local habitat conditions promoting the presence of the host and vector species in the Kazakh plague system. *PLoS ONE*, 10(9), p.e0136962.

Laundré, J.W. & Reynolds, T.D., 1987. Effects of soil structure on burrow characteristics of five small mammal species. *Great Basin Naturalist*, 53(4), pp.358–366.

Morse, S.S., 1995. Factors in the emergence of infectious diseases. *Emerging infectious diseases*, 1(1), pp.7–15.

Naumov, N.P. & Lobachev, V.S., 1975. Ecology of desert rodents of the USSR (Jerboas & Gerbils). In I. Prakash & P. K. Ghosh, eds. *Rodents in Desert Environments*. The Hague: The Hague, pp. 465–598.

Onishchenko, G.G. & Kutyrev, V.V. eds., 2004. *Natural plague foci in the Caucasus, Caspian Sea region, Middle Asia and Siberia.*, Meditsina Publishers (In Russian).

Pogrebinsky, M.A., 1963. Illi valley, its nature and resources. *Soils of Kazakhstan*, 4(422 (In Russian)), pp.306–326.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. Available at: http://www.r-project.org/.

Randall, J.A. et al., 2005. Flexible social structure of a desert rodent, Rhombomys opimus: philopatry, kinship, and ecological constraints. *Behavioral Ecology*, 16(6), pp.961–973.

Rapoport, L.P. et al., 2010. Comparative analysis of the flea fauna and its epizootic importance in the deserts of Southern Kazakhstan. *Entomological Review*, 90(8), pp.1003–1013.

Reichman, O.J. & Smith, S.C., 1990. Burrows and Burrowing Behaviour by Mammals. In H. H. Genoways, ed. *Current Mammalology*. New York & Longon: Plenum Press, pp. 369–416.

Reijniers, J. et al., 2012. A curve of thresholds governs plague epizootics in Central Asia. *Ecology letters*, 15(6), pp.554–60.

Rhodes, D.H. et al., 1985. Influence of Soil Texture , Moisture and Temperature on Nest-site Selection and Burrowing by the Pine Vole , Microtus pinetorum. *American Midland Naturalist*, 113(1), pp.102–108.

Rivkus, J.Z. et al., 1992. *Epidemiology and prophylaxis of the plague.*, Ashkhabat: In Russian.

Roberts, H.R. et al., 2014. Southwestern Association of Naturalists Burrowing Ecology of Pocket Gophers (Rodentia : Geomyidae) in Jalisco , Mexico. *The Southwestern Naturalist*, 42(3), pp.323–327.

Schmid, B. V. et al., 2015. The pulse of Asia: the Black Death and successive reintroductions of plague into Europe. *Proceedings of the National Academy of Sciences of the United States of America*, Early Edit, pp.1–6.

Semachko, L.L., 1981. House and tree sparrows as carriers of ticks in Turkmenia,. *Zod. Zh*, 40(7), pp.1070–1078.

Shenbrot, G. et al., 2002. Habitat-dependent differences in architecture and microclimate of the burrows of Sundevall's jird (Meriones crassus) (Rodentia: Gerbillinae) in the Negev Desert, Israel. *Journal of Arid Environments*, 51(2), pp.265–279.

Stenseth, N.C. et al., 2008. Plague: past, present, and future. *PLoS medicine*, 5(1), p.e3.

Stenseth, N.C. et al., 2006. Plague dynamics are driven by climate variation. *Proceedings of the National Academy of Sciences of the United States of America*, 103(35), pp.13110–5.

Telfer, S., 2010. Species Interations in a Parasite Community Drive Infection Risk in Wildlife Population. *Science*, 330, pp.243–246.

Venables, W.N. & Ripley, B.D., 2002. *Modern applied statistics with S.* Fourth., New York: Springer.

Wentworth, C.K., 1992. A Scale of Grade and Class Terms for Clastic Sediments. *The Journal of Geology*, 30(5), pp.377–392. Available at: http://www.jstor.org/stable/30063207.

Whitford, W.G. & Kay, F.R., 1999. Biopedturbation by mammals in deserts: a review. *Journal of Arid Environments*, 41(2), pp.203–230.

Wikel, S.K., 2011. Modulation of the Host Immune System by Ectoparasitic Arthropods. *Bioscience*, 49, pp.311–320.

Wilkins, K.T. & Roberts, H.R., 2007. Comparative analysis of burrow systems of seven species of pocket gophers (Rodentia: Geomyidae). *The Southwestern Naturalist*, 52(1), pp.83–88.

Wilschut, L.I. et al., 2013. Mapping the Distribution of the main host for plague in a complex lanscape in Kazakhstan : An object-based approach using SPOT-5 XS , Landsat 7 ETM + , SRTM and multiple Random Forests. *International Journal of Applied Earth Observation and Geoinformation*, 23, pp.81–94.

Wilschut, L.I. et al., 2013. Potential corridors and barriers for plague spread in Central Asia. *International journal of health geographics*, 12(1), p.49.

Xu, W. et al., 2012. Rhombomys opimus contribution to the “fertile island” effect of tamarisk mounds in Junggar Basin. *Ecological Research*, 27(4), pp.775–781.

3. Predicting gerbil occupancy: a metapopulation model, and assessment of population connectivity.

3.1. Introduction

The great gerbil, *Rhombomys opimus* (Licht 1823) is the dominant reservoir host species of the gram-negative bacterium *Yersinia pestis* in the pre-Balkhash desert located to the North East of Kazakhstan. Whilst infections have been observed in other mammal species in the area such as the midday gerbil (*Meriones spp.*), only *R.opimus* populations are observed to have a consistent prevalence of infection over long periods of time (Begon et al. 2006; Davis et al. 2007a), and as such are considered a lone sylvatic reservoir species in this area.

Previous theoretical work on the system has demonstrated that periods of high gerbil infection prevalence, referred to as epizootics, are well predicted by the density of *R. opimus* individuals in a given area (Davis et al. 2007b; Davis et al. 2008). Specifically, the epizootic is only able to spread across a gerbil population that is sufficiently connected through series of occupied burrows lying in close proximity. The spread of *Y. pestis* then acts as a percolation phenomenon across the burrows (Davis et al. 2008). Predictions of outbreaks can be improved further by the inclusion of estimates of flea burden per animal, but this still relies upon the underlying relationship with gerbil connectivity (Reijniers et al. 2012). As such, understanding and being able to make predictions about gerbil population size and spatial structure are fundamental to predicting *Y. pestis* outbreaks. The gerbil population is spatially structured by virtue of their social behaviour, with a family unit typically comprising of one male, one or two females and one or two young remaining associated with one burrow structure over several seasons. One family may potentially cycle through multiple generations at the same burrow, with some younger individuals moving on to establish new family groups in other burrows (Randall et al. 2005). These burrows are typically long-lasting, with complex structure and distinct compartments, that individuals more frequently move into rather than constructing new burrows (Naumov & Lobachev 1975).

Current models of the Kazakh plague system appear to over-estimate the frequency of *Y. pestis* epizootic events in the area (Reijniers et al. 2012; Reijniers et al. 2014). One hypothesis to explain this is that the current measures of gerbil presence, given as a density typically over a trapping sector (10x10km area) or gerbils

per km^2, may not accurately reflect processes occurring at a closer range. In other words; does the movement of gerbils between specific burrows alter their population size and structure, and as a consequence the dynamics of *Y. pestis* epizootics, in a manner that cannot be observed by only considering their numbers over a larger area? Testing this is complicated by the gerbils' choice of habitat, as the large part of their life that is subterranean has made recording their movement behaviours, and estimating related parameters such as habitat range or dispersal rate difficult. Studies have been conducted using GPS collar tracking, but data is limited and signals often become interrupted once gerbils return to their burrows and move a certain distance below ground (Hughes et al. 2015).

Separate studies have identified that satellite images can be used to predict burrow occupancy (Wilschut et al. 2013). Given great gerbil burrows are typically at least semi-permanent structures within the landscape (Naumov & Lobachev 1975; Begon et al. 2006) a modelling approach describing the colonisation behaviour of the gerbils combined with satellite images describing current burrow position and occupancy would represent a potentially powerful strategy for predicting gerbil population fluctuations without large and generally costly fieldwork expeditions.

Spatial heterogeneity and the natural subdivision of populations is observed in a variety of natural species (Hanski & Gilpin 1991). Describing and predicting the behaviour of such populations has therefore remained a common challenge in many areas of ecology (Alexander et al. 2012). Referred to as "metapopulations", mathematical models can be used to simulate such systems as a series of patches, which at a given time are considered to be occupied (the species of interest is present) or unoccupied (the patch is currently not used by the species of interest). The first mathematical model of a metapopulation was proposed by Levins in 1969 (Levins 1969), who defined a system with a constant rate of patch extinction, and a rate of patch colonisation proportional to the number of patches occupied at a given time. Whilst this captures many of the key dynamics of a metapopulation, the Levins model ignores the spatial structure of the patches, effectively allowing for colonisation events at a patch to arise from individuals moving from any other patch. Given behavioural or simply physical restrictions on movement in many species, this will typically not reflect the real population dynamics. Developments of this model to allow for animal movement across the landscape, including Spatial Patch Occupancy Models (SPOMs),

take these underlying properties of the Levin's model but explicitly consider the patch level dynamics, and allow for the stochasticity observed in these systems to be accounted for (Hanski 1991; Hanski 1994; Hanski & Ovaskainen 2003). Similarly, extinction events occur at a constant rate, but at each time step the rate of colonisation at a given patch depends on the number of occupied patches within a given proximity.

The speed and efficiency of plague transmission within individuals inhabiting the same burrow is such that we can consider the presence of one infected animal within a burrow to to be representative of any eventual infection and moreover onward transmission arising from the burrow (Begon et al. 2006; Davis et al. 2007a; Davis et al. 2007b). For the purposes of monitoring and projecting the gerbil population in order to make predictions about *Y. pestis* infections, the population can therefore be considered within a metapopulation framework with burrows as patches. In the previous Chapter, it was investigated whether environmental properties associated with individual burrows were predictive of resulting gerbil or flea presence. This hypothesis was not rejected, with properties such as the presence of sand dunes, trees, and alluvial sediments being significantly associated with probability of gerbil presence at a burrow site (Levick et al. 2015). A spatially explicit model would permit this patch hetereogeneity to be integrated into the model.

Metapopulation approaches have been utilised to model plague dynamics across a range of systems. Early approaches focussed on anticipating human dynamics rather than being specific to a particular wild focus (Keeling & Gilligan 2000). Even in this early model, parameterised in a much less system specific manner than its descendants, this identified key aspects of *Y. pestis* epidemiology including the divergent patterns of outbreaks and persistent endemicity amongst animal populations. Subsequently, metapopulation techniques were utilised in the previous studies of *R. opimus* threshold dynamics in Kazakhstan discussed above (Davis et al. 2007b).

In epidemiology, a critical community size (CCS) is a minimum susceptible population size required for an infectious agent to persist, for example as demonstrated for measles in human populations (Bolker & Grenfell 1995). Hanski & Ovaskainen (Hanski & Ovaskainen 2003) identified that in spatially structured populations, when the total population sizes falls below the CCS, it would be expected that spatial fragmentation of the resulting occupied patches would favour any

infectious agent to which the population plays host. This assertion was challenged, for the persistence of *Y. pestis* in Kazakhstan in periods of low *R. opimus* numbers (Jesse & Heesterbeek 2011). In particular, Jesse and Heserbeek identified that persistence is related non-linearly to movement distance (which itself may be underestimated, by failure to consider the direction and nature of movements). Moreover, they show that movement, epidemiological properties and host demography all contribute to persistence. Considering the importance of these other drivers, they argue that the relationship between *Y. pestis* persistence and *R. opimus* population fragmentation is not unequivocally demonstrated. Therefore, we may expect that models of the *Y. pestis* system in Kazakhstan may diverge from expected metapopulation behaviour.

A number of metapopulation studies of the *Y. pestis* system in Kazakhstan aims related to the identification of, or theoretical assessment for the existence of small sub regions able to sustain the bacterium, referred to as hotspots (Davis et al. 2007a). These hotspots are hypothesised to act as refugia for *Y. pestis* in the absence of sufficient host numbers. However, recent theoretical work suggests that such hotspots would need to be large areas with sustained gerbil population densities, unlike those observed empirically in Kazakhstan (Schmid et al. 2012). This was corroborated by a more recent study using an Agent Based Model (ABM) to inform a deterministic compartmental population based model (PBM). The ABM integrated a great deal of detail on both gerbil and flea population dynamics and demography, to be implemented in the PBM. Even this more specifically informed approach failed to identify evidence for hotspots (Laperrière et al. 2016).

Generally, simulation and modelling approaches of *R. opimus* in Kazakhstan utilise a random distribution or lattice of burrows, reflective of known large scale measures such as gerbil abundance and burrows per hectare (Davis et al. 2008; Reijniers et al. 2012; Jesse & Heesterbeek 2011; Schmid et al. 2012; Laperrière et al. 2016). The present study is able to inform the model with population records, spatial locations and ecological properties of real individual burrows. Previous statistical analysis of the distribution of burrows and occupancy identified the role of spatial structures in driving gerbil occupancy (Wilschut et al. 2013), and that clustering is observed throughout the system (Wilschut et al. 2015). The latter in particular discusses the potential that the patch in the *Y.pestis* metapopulation in Kazakhstan may include multiple burrows, although it could not be definitively identified how

many burrows this may include (or alternatively, what area of space it would cover). This investigation is able to consider these questions again in an explicit manner, including greater detail of individual burrows, and by generating simulations of the population. In particular, given that in Chapter 2 it was identified that individual burrow properties were related to their gerbil dynamics, it may be that through considering this we are able to identify more localised patterns of population behaviour. As such we are able to test hypotheses about the role of different drivers in the gerbil population dynamics. Ultimately then, the question remains how unknown factors in the gerbil population dynamics may drive, and be used to inform the dynamics of *Y. pestis* infection. Here the question is posed of whether this can be answered by characterising burrow level effects.

This investigation identifies that the standard SPOM approach was unable to generate simulated results approaching the empirical observations of the gerbil population, even with real burrow locations and allowances for seasonal dynamics. As an alternative, the connectivity of individual burrows was estimated, although again this was found to not fully explain population dynamics. The power of this second relationship was explored through the generation of simulated data sets. This analysis suggests that the failure to identify relationships may be a power issue. This technique is explored in elsewhere (Levick, Begon and Cornell, in prep.). Given the previous studies discussed above, we conclude that a good model of gerbil population dynamics could be highly informative, but that a combination of further field work and an exploration of patch size should be considered going ahead.

3.2. Aims and hypotheses

Aim

- Generate projections of the *R. opimus* population incorporating dynamics at the individual burrow scale.

Hypotheses

- Occupancy at an individual burrow at a given time is determined by the location of surrounding occupied burrows.
- Any relationship between burrow occupancy and the location of surrounding occupied burrows may vary with:
    - Environmental properties of the burrow,

- The season in which the burrow is observed.
- Occupancy at an individual burrow at a given time is determined by the location of surrounding burrows alone, irrespective of their occupancy status.

3.3. Methods

Field data

A set of 513 burrows were selected from across 5 of sectors (10x10km trapping areas) across the pre-Balkhash focus and their position recorded as Latitude/Longitude coordinates using a handheld GPS device. The burrows were visited over 7 trapping seasons (spring, summer, and autumn of 2011 and 2012, and spring of 2013). On each occasion, the occupancy status of the burrow was recorded. The presence of faeces or clear digging marks would indicate that the burrow was occupied. Feet and tail markings near the entrance would indicate that animals had visited, but the burrow was not currently occupied. For the purposes of this analysis these burrows were considered unoccupied. Burrows absent of any indicators of gerbil presence would be classed as unoccupied.

Metapopulation model

Based on a SPOM approach, a model was built to simulate the gerbil population as a metapopulation with individual burrows as patches.

The model is controlled by three parameters; extinction rate $\boldsymbol{\mu}$, colonisation frequency $\boldsymbol{C_0}$ and connectivity range $\boldsymbol{L}$. The latter two are used in simulations to calculate the colonisation rate $\boldsymbol{K_i}$ for each burrow $\boldsymbol{i} \in \boldsymbol{u}$ (where $\boldsymbol{u}$ is the set of unoccupied burrows), given by equation 3.1.

$$\boldsymbol{K_i = C_0 \sum_{j \in o} e^{\frac{-|r_i - r_j|}{L}}} \qquad 3.1$$

Where **L** is the colonisation range estimated for all burrows as described below, $\mathbf{r_i}$ & $\mathbf{r_j}$ the real positions of burrows at positions $\boldsymbol{i}$ and $\boldsymbol{j}$ (given as latitude-longitude coordinates) where $\boldsymbol{j \in o}$, (where $\boldsymbol{o}$ is the set of occupied burrows), and $|\mathbf{r_i} - \mathbf{r_j}|$ the absolute distance between them.

As such the colonisation rate at an individual burrow will be higher the more burrows there are located nearby.

This value is then converted to a colonisation probability $\boldsymbol{C_i}$, bound between 0 and 1, using equation 3.2.

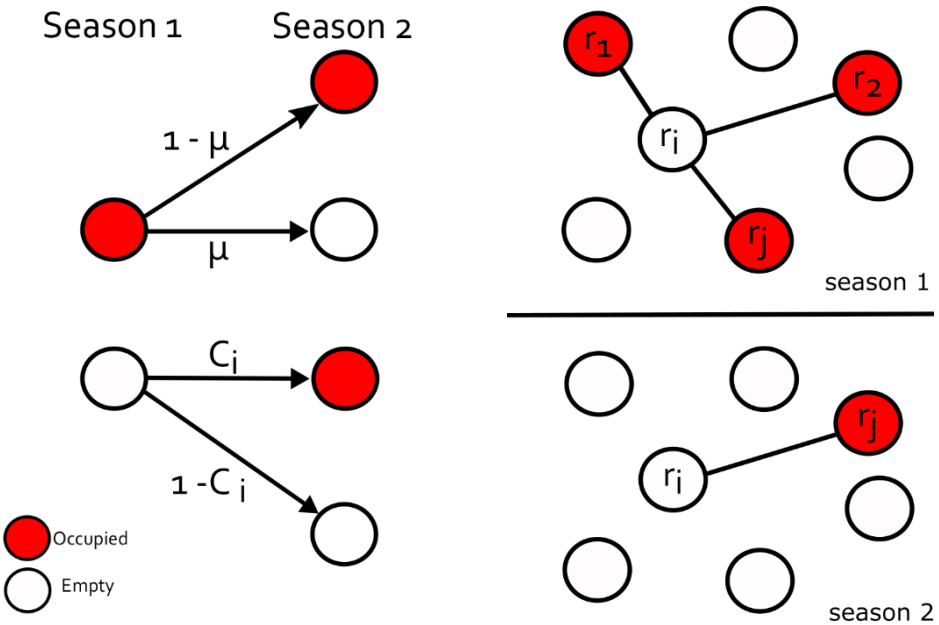
$$\boldsymbol{C_i} = \frac{\boldsymbol{K_i}}{\mathbf{1} + \boldsymbol{K_i}} \qquad 3.2$$

Where $\boldsymbol{K_i}$ is the value calculated for each burrow using equation 3.1.

This colonisation probability and the extinction rate then drive simulations of the system as detailed below.

Simulations with a single colonisation frequency estimate

For each season a state vector was constructed with length 513 (a position for each burrow), containing a 1 for each occupied burrow and 0 for an unoccupied burrow. In every case, the first vector was a vector of observed burrow occupancy records from the data, referred to as the seed data. For each subsequent time step a new vector would be constructed representing the occupancy status for each burrow in that generation. If a burrow was occupied at the starting time point, it would become extinct with the probability $\boldsymbol{\mu}$, and conversely remain occupied with probability $\mathbf{1} - \boldsymbol{\mu}$. Similarly, an unoccupied burrow $\boldsymbol{i}$ would become occupied with the probability $\boldsymbol{C_i}$ (Figure 3.1A). The colonisation probability $\boldsymbol{C_i}$ would be obtained by calculating the colonisation parameter $\boldsymbol{K_i}$ at each unoccupied burrow $\boldsymbol{i} \in \boldsymbol{u}$, where $\boldsymbol{u}$ is the total unoccupied burrows in starting season at each time step, using the estimated values of $\boldsymbol{C_0}$ and $\boldsymbol{L}$. As such this would reflect the burrows proximity to burrows occupied at the current time step (Figure 3.1B).



(A) *(B)*

Figure 3.1. (A) Simulation process per burrow. Depending on the starting status (season 1) the burrow would either stay the same or change, at the given probabilities. (B) The colonisation probability is calculated at each time step to reflect the current spatial distribution of occupied burrows.

Simulations with seasonally variant colonisation frequency estimates

The gerbil population does not grow at a constant rate throughout the year, rather there is a breeding season in spring, leading to high numbers in summer and autumn, followed by a die off over winter (Naumov & Lobachev 1975). To reflect this, a second simulation was conducted with variant values for the baseline connectivity parameter $\boldsymbol{C_0}$ depending on the season transition at that time point. For example, assuming the first set of observations are in Spring, the parameter describing spring to summer colonisation parameter ($\boldsymbol{C_{sp,su}}$) would be used to calculate $\boldsymbol{K_i}$ as shown in equation 3.3.

$$\boldsymbol{K_i = C_{s1,s2} \sum_{j \in o} e^{\frac{-|r_i - r_j|}{L}}} \qquad 3.3$$

Where s_1 and s_2 are the start and end season of the, e.g. s_{sp}, s_{su} showing a spring to summer transition. Three forms of this existed for spring to summer, summer

to autumn and autumn to spring transitions respectively. Transitions were considered from autumn to spring as data could not be collected in the winter, so the parameter could only be estimated for an autumn to spring transition. Equally this is reflected in the simulations, where each time point would be either spring, summer or autumn.

In both cases, simulations were run for a total of 7 seasons (6 transitions) and iterated 500 times.

Parameter estimation

Parameter estimates at maximum likelihood (using a computed binomial likelihood), given the data were estimated for $\boldsymbol{C_0}$ and $\boldsymbol{L}$ given the functional form to calculate $\boldsymbol{K_i}$ as shown in equation 3.1. , , Here $\boldsymbol{i} \in \boldsymbol{n}$, where $\boldsymbol{n}$ is the set of all burrows and $\boldsymbol{j} \in \boldsymbol{o}$ where $\boldsymbol{o}$ is the set of occupied burrows at a time step. The extinction rate $\boldsymbol{\mu}$ was estimated from the number of successful extinction events in the data. These were both performed via numerical optimisation, using the Brent method to estimate $\boldsymbol{\mu}$ and Nelder-Mead to estimate $\boldsymbol{C_0}$ and $\boldsymbol{L}$.

Seasonally variant values of the colonisation parameter were estimated similarly but from relevant subsets of the data, i.e. $\boldsymbol{C_{0,spsu}}$ would be estimated using the above procedure but only using data where the starting season were from spring records and the second season from summer records.

Statistical modelling of environmental properties

A previous study suggested that environmental heterogeneities at the level of the individual burrow are significantly associated with the probability that a burrow would become, and stay occupied (Levick et al. 2015). To identify whether similar associations could be identified with the probability of either colonisation or extinction at the burrow specifically (rather than just a general measure of occupancy), binomial logistic regression models were constructed. Details of the factors included are given in Table 3.1. As in the previous study, the trapping sector to which each burrow belonged was included in the models as a random effect, and the models implemented using the GLMM functions from the MASS package (Venables & Ripley 2002). To avoid over-complication models were constructed for terms relating to the sediment properties (sand, clay, etc.) and to environmental properties (latitude position, presence of trees) first. Then factors found to have a significant association

were brought together to form a maximum model, which would then be reduced via stepwise selection to a minimal model.

Table 3.1. Descriptions of environmental factors describing burrows included in GLMM's to find associations with extinction and colonisation events.

Model Variable	Description
Sand	Binary, presence or absence of sand
Clay	Binary, presence or absence of clay
Solonchak/takir	Binary, presence or absence of solonchak/takir sediment
Dunes	Binary, presence or absence of sand dunes
Tree	Binary, presence of absence of at least 1 tree at the burrow site
Latitude	Latitude position of burrow in °N

Landscape Connectivity

The simulations described above rely on parameters that require at least two seasons of occupancy data at a given burrow to be included in the parameter fitting. Some of the burrows did not have a complete set of observations. Furthermore, the observations were limited to the study window. The data were limited in general in both breadth and coverage, and this is a consistent challenge due to the difficulties associated with collecting field data in this system. In terms of future investigations, it may be useful to know if the spatial distribution of the burrows alone, ignoring occupancy status at a given moment, is sufficient to predict colonisation probability at a burrow. From their spatial location and relationship to the distribution of other burrows we can instead estimate landscape connectivity (Moilanen & Nieminen 2002) and identify whether this is sufficient to predict colonisation events.

The landscape connectivity $\boldsymbol{h_i}$ is then calculated using a modified form of equation 3.1 shown in equation 3.4.

$$\boldsymbol{h_i} = \sum_{\boldsymbol{j} \in \boldsymbol{n}} \boldsymbol{e}^{\frac{-|r_i - r_j|}{\boldsymbol{b}}} \qquad 3.4$$

Where both $\boldsymbol{i}$ and $\boldsymbol{j} \in \boldsymbol{n}$, where $\boldsymbol{n}$ is the set of all burrows. The parameter $\boldsymbol{L}$, connectivity range used previously is replaced by the parameter $\boldsymbol{b}$ describing maximum gerbil movement range. This effectively weights the contribution of a pair of patches: well-connected patches have a higher landscape connectivity. Where before the distance between the burrows $|\mathbf{r_i} - \mathbf{r_j}|$ was calculated as the Euclidean distance between the coordinates, here this was converted to a distance in kilometres, as this parameter can be considered to reflect the real distance over which gerbil movement is likely to occur.

An estimate for gerbil movement range, $\boldsymbol{b}$ at maximum likelihood given the data was again identified using numerical optimisation. In this instance, corresponding connectivity values were calculated for each burrow at the current value of $\boldsymbol{b}$, and a logistic regression model constructed of associations of the values with either colonisation or extinction events. This was then optimised to minimise the log likelihood of the resulting model.

Statistical modelling of landscape connectivity

To identify whether connectivity alone was responsible for colonisation or extinction events, the occupancy data was converted to colonisation opportunities and successes for each season transition at each burrow. The resulting matrices were used to conduct binomial logistic regression analysis using connectivity values calculated using the estimated value of $\boldsymbol{b}$, and the occupancy levels of the surrounding sector as predictors. The trapping sector was specified as a random effect.

The theoretical limits of such a relationship were explored by identifying the minimum effect size required for connectivity to be a significant predictor of occupancy using a modified power calculation. Iterating through different values of the coefficient for the connectivity term, at each a new set of response data was

generated based on the original GLMM. Estimates for the intercept value were iterated over, and one chosen where the total proportion of possible colonisation/extinction events that were successful in the predicted data matched that of the empirical dataset. This new data was then used as a response data in a new GLMM to identify the resulting effect size and significance. These GLMMs were constructed using the glmmADMB function from the package of the same name (Skaug et al. 2013; Fournier et al. 2012), as this function allows for the calculation of logLikelihoods and associated values such as Akaike's Information Criterion (AIC) scores.

To explore the utility of this approach where multiple season occupancy data is unavailable, this process was then repeated, using a larger data set generated using the original model and GPS point coordinates for an addition 3514 burrows. As discussed in Chapter 2, field data was collected from a larger selection of burrows over the same time period and geographical area, but detailing only the burrows' GPS position and occupancy status in one, or less frequently two seasons. To generate multiple seasons' worth of data, each burrow was assigned a number of possible colonisation events as either 3 or 4, with each occurring with a probability matching their frequency in the long term data set. These were used along with the model to repeat the process above to generate response data, and similarly the association between connectivity and colonisation success measured at different coefficient values. The implemented method is discussed in greater detail elsewhere (Levick, Begon and Cornell, in prep.).

All analysis was conducted using the R Statistical Computing package (R Core Team 2014), using the MASS (Venables & Ripley 2002) and glmmADMB (Fournier et al. 2012; Skaug et al. 2013) packages.

3.4. Results

Metapopulation Simulation

Parameter estimation: single colonisation frequency

Assuming a constant rate of colonisation in each possible season transition, values of the constant parameters $\boldsymbol{C_0}$, $\boldsymbol{L}$ and $\boldsymbol{\mu}$ were estimated from the data at maximum likelihood given the functional form for the colonisation parameter given in equation 3.1 and are detailed in table 3.2.,

Table 3.2. Parameter estimates and associated negative log likelihood scores for the baseline colonisation rate and range, and the extinction rate.

Parameter	Estimate	-logLik	Optimizer
Baseline colonisation rate, C_0	0.007	815.640	Nelder-mead
Baseline colonisation range, L	0.023		
Extinction rate, μ	0.328	609.836	Brent

Simulations: single colonisation frequency

Simulations were then performed, running for 7 generations of occupancy status for each burrow (as in the field data), iterated over 500 realisations of the simulation. The results are shown below in Figure 3.2. The simulated occupancy records do not appear to follow the same behaviour, or sit around the same proportion of burrows occupied, as the field data sets.

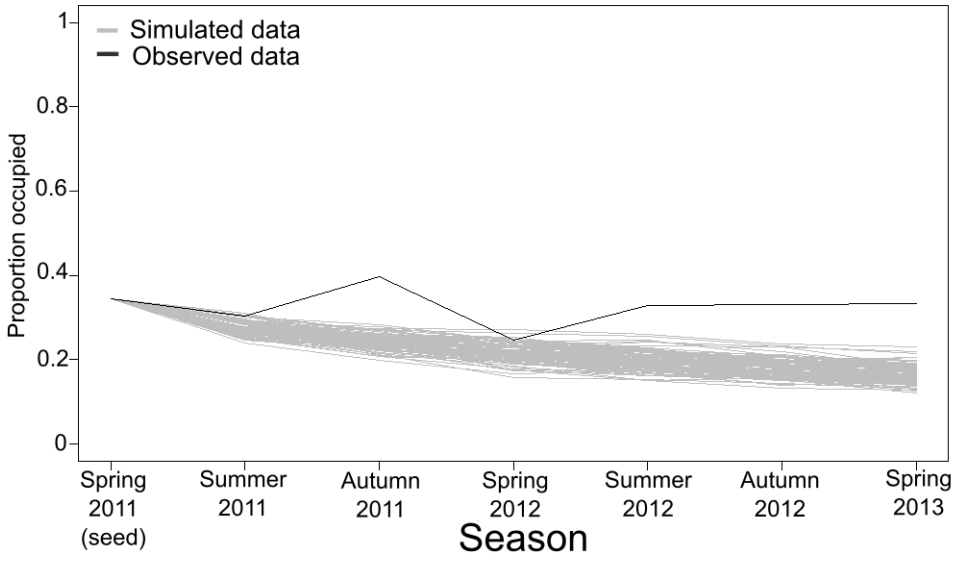


Figure 3.2. Output from simulations assuming a constant baseline colonisation frequency in any given season transition. The data at each generation is summarised as the proportion of burrows occupied that season. The black line indicates the observed field data, and the grey lines each of the realisations of the simulations.

Parameter estimation: seasonally variant colonisation frequencies

Whilst the estimate for $\boldsymbol{\mu}$ remained the same, the colonisation range was re estimated and seasonally varying colonisation frequencies estimated using subsets of the data at maximum likelihood using equation 3.3, and are detailed in Table 3.3.

Table 3.3. Parameter estimates and associated negative log Likelihood scores for the colonisation range, and baseline colonisation rates for each possible two season transition type (spring-summer, summer-autumn, autumn-spring).

Parameter	Estimate	-logLik	Optimiser
Baseline colonisation range, $\boldsymbol{L}$	0.010		
Baseline colonisation rate, $\boldsymbol{C_{spsu}}$	0.015	810.005	Nelder-Mead
Baseline colonisation rate, $\boldsymbol{C_{suau}}$	0.011		
Baseline colonisation rate, $\boldsymbol{C_{ausp}}$	0.009		

Simulations: seasonally variant colonisation frequencies

These parameters were used to run further simulations of the gerbil population over 7 seasons, through 500 realisations. The results of these simulations are summarised in Figure 3.3. The behaviour of the occupancy records is once more not reflected in the simulated data, although the proportion of burrows occupied appears to move closer to that in the empirical data.

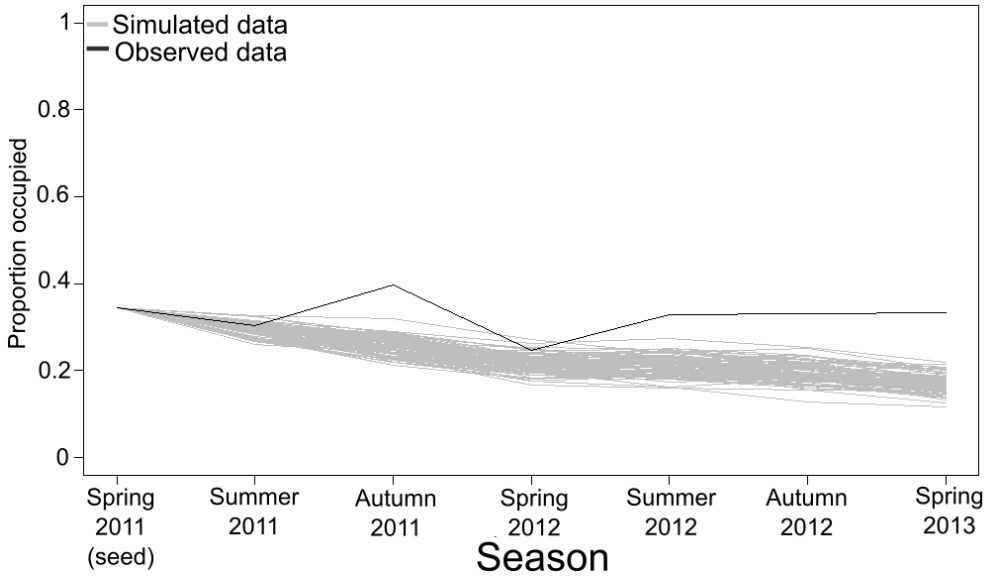


Figure 3.3. Output from simulations with baseline colonisation rates adjusted according to the particular seasons at the start and end of the colonisation event. The data at each generation is summarised as the proportion of burrows occupied that season. The black line indicates the observed field data, and the grey lines each of the realisations of the simulations.

Statistical modelling of environmental properties

For both extinction and colonisation events, only the latitude position of the burrow was found to be significantly associated (in both cases $P<0.001$, and all other $P>0.05$). The results are detailed in Tables 3.4 and 3.5 below. These results therefore differed to those obtained previously for the same predictors considered for burrow occupancy and long term occupancy.

Table 3.4. GLMM output of sediment and large scale environmental variables associated with colonisation events.

Factor	Coefficient	Standard error	P Value	Model
Presence of sand	0.56	0.32	0.09	
Presence of dunes	-0.11	0.26	0.68	
Presence of clay	0.12	0.27	0.66	Sediment
Presence of solonchak/takir	0.69	0.76	0.36	
(Intercept)	-2.07	0.54	<0.001	
Latitudinal position	3.58	29.58	<0.001	
Presence of at least 1 tree	-0.06	0.20	0.74	Environment
(Intercept)	-163.86	29.58	<0.001	

Table 3.5. GLMM output of sediment and large scale environmental variables associated with extinction events.

Factor	**Coefficient**	**Standard error**	**P Value**	**Model**
Presence of sand	-0.13	0.37	0.72	
Presence of dunes	-0.02	0.27	0.93	
Presence of clay	0.01	0.32	0.96	
Presence of solonchak/takir	-0.27	0.72	0.70	Sediment
(Intercept)	-0.20	0.45	0.66	
Latitudinal position	-2.25	0.46	<0.001	
Presence of at least 1 tree	-0.23	0.18	0.21	Environment
(Intercept)	102.15	20.99	<0.001	

Parameter estimation: Landscape Connectivity

The connectivity range parameter ***b*** was estimated for colonisation probability given the landscape connectivity values calculated using equation 3.4. The estimate and log likelihood is given in Table 3.6.

Table 3.6. Connectivity range parameter ***b*** estimate.

Model	Value of ***b***	-logLik
Colonisation	0.11	421.77

Statistical modelling of landscape connectivity

From the estimate of $\boldsymbol{b}$, values of landscape connectivity $\boldsymbol{h_i}$ were calculated for each burrow. Generalized linear mixed binomial models identified that given the average occupancy of the surrounding sector, the connectivity of the individual burrow was not significantly associated with either colonisation or extinction success. Sector occupancy was however significantly associated with colonisation success (P<0.001) (Table 3.7.)

Table 3.7. GLMM output of occupancy and connectivity estimates associated with colonisation and extinction probability.

Response	**Factor**	**Coefficient**	**Standard error**	**P Value**
Colonisation success	Connectivity	0.038	0.028	0.18
	Sector Occupancy (gerbils per km2)	79.76	12.98	<0.001
	(Intercept)	-3.87	0.37	<0.001

Simulation models to explore effect size and power

Iterating through potential coefficient sizes for the connectivity term in the model of colonisation events identified that even at a high value of the coefficient, a significant relationship could not be observed (Figure 3.4).

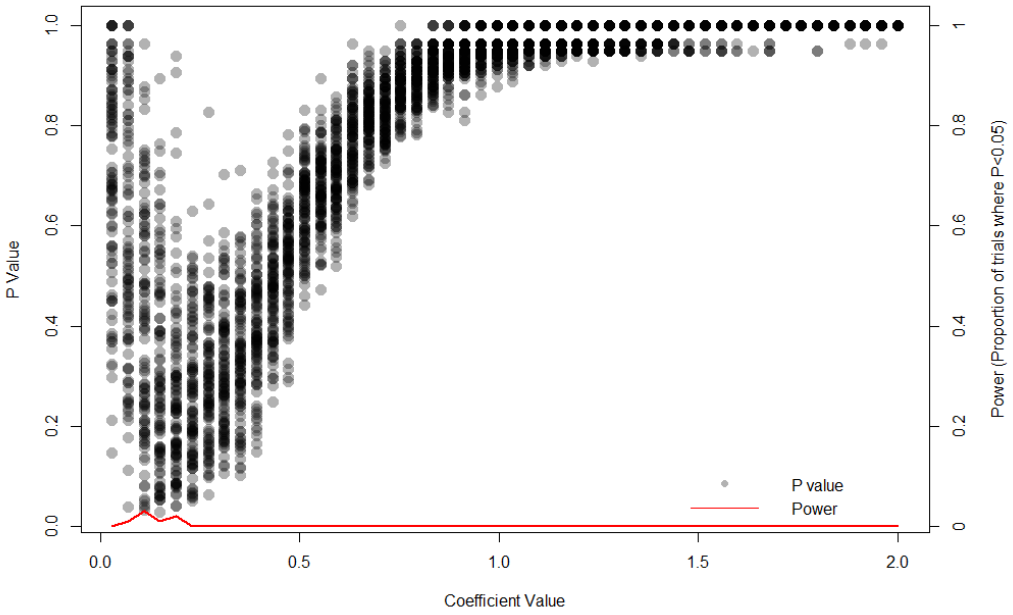


Figure 3.4. P values from simulated data for generalized linear mixed models identifying associations between a burrow's connectivity and the occupancy of its surrounding sector with its colonisation success, where the sector is set as a random effect. The model was run for 100 realisations at 50 coefficient values for the connectivity term, using simulated data generated from the fitted model with the new coefficient value.

However, iterating across potential coefficient values in the expanded data set showed that from weak coefficients to a value of around 0.9, a significant association between connectivity and colonisation success could be observed. This was not observed in the upper ranges of the coefficients tested (Figure 3.5).

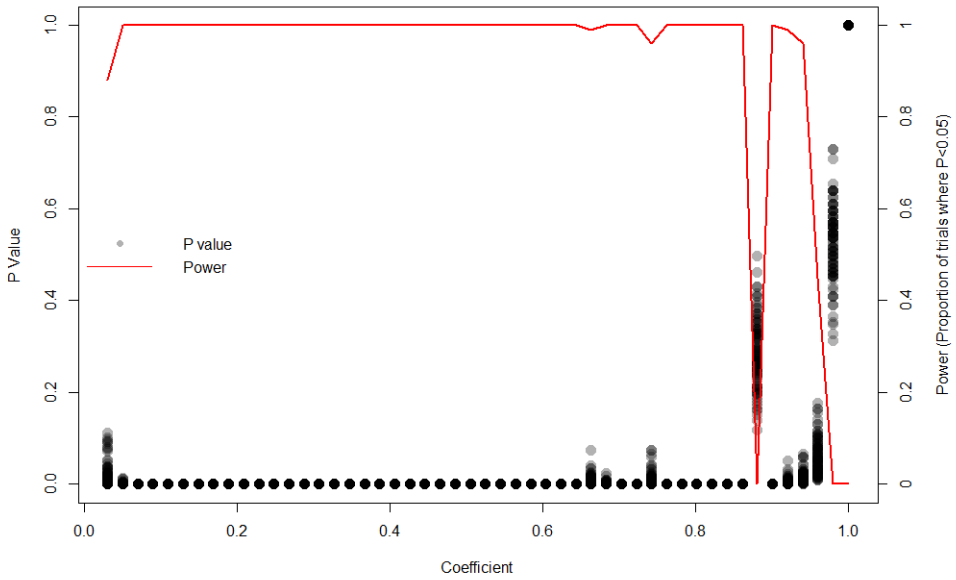


Fig 3.5. P values from simulated data for generalized linear mixed models identifying associations between a burrow's connectivity and the occupancy of its surrounding sector with its colonisation success, where the sector is set as a random effect. This model was run for 100 realisations at 50 values for the coefficient of the connectivity term, using simulated data generated from the fitted model with the new coefficient value, using additional burrows whose long term occupancy is unknown, but their connectivity (h) values could be estimated from their GPS position and those surrounding it.

3.5. Discussion

Previous theoretical and empirical work concerning *Y. pestis* in wild *R. opimus* gerbils in the pre Balkhash desert has revealed that the structure and size of the gerbil population is a central driver to *Y. pestis* transmission and prevalence (Davis et al. 2004; Begon et al. 2006; Davis et al. 2007a; Reijniers et al. 2012). However, even the most recent models typically predict that plague will be present in areas where it is then subsequently not observed (Reijniers et al. 2012), driving further research to improve these projections. Metapopulation approaches have been utilised to investigate *Y. pestis* infection dynamics in Kazakhstan, demonstrating key epidemiological properties of the system; that it may not align with previous expectations of fragmentation in such systems (Jesse & Heesterbeek 2011) and that hotspots alone may not be sufficient to explain periods of undetectably low prevalence (Schmid et al. 2012). However, whilst it has been shown that the gerbil population

dynamics determine *Y.* pestis dynamics, models of the epizootics rely upon large scale measures to generate simulated networks of burrows. The most detailed approach was taken by Lapierre *et al.* (Laperrière et al. 2016), however the present study is the first to generate simulations of this population informed by ecological and population data collected at the scale of the individual burrow.

The spatial distribution of the gerbil population has been shown to be influenced by environmental structures, forming corridors and barriers for the population (Wilschut et al. 2013), and clustering of burrows is observed (Wilschut et al. 2015). Further, it has been demonstrated that local burrow conditions are associated with the probability of a burrow being, or remaining occupied (Levick et al. 2015). The present study aimed to take this close resolution approach further, and investigate whether burrow specific properties such as real burrow geolocations would impact projections of overall population dynamics. It was demonstrated that the population is probably not well reflected by the metapopulation model as constructed here. The proximity of each burrow to those around it (or connectivity) may determine the probability of burrows becoming occupied. However, this association could only be observed in a much larger sample, of data generated using the field records as a training data set. As such the connectivity of a burrow to those around it may be a useful predictor for gerbil, and therefore *Y. pestis* presence, but a much larger study would be required to define this relationship empirically.

An association between connectivity and extinction was also not observed, but this is less surprising, given extinction is more likely determined by environmental or within burrow aspects (Randall et al. 2005; Naumov & Lobachev 1975). What is perhaps not considered fully here is the role of burrows that, at least as far as our data can inform us, do not change occupancy status and remain either occupied or empty for the entire series. For example, in the colonisation data, a burrow that remains occupied for the entire series will not be included, because an opportunity for colonisation was not observed. However, the work discussed in Chapter two suggests that this may be a high quality burrow (Levick et al. 2015). It may be the case that these burrows are losing their family groups, but are becoming reoccupied before the next season so we fail to observe this. Further study is beginning in the area concerning gerbil movements and population dynamics (Yeszhanov, Pers. Comms.) that may be able to offer more insight.

R. opimus burrows are distinct structures in the landscape of the pre-Balkhash, and both their location and occupancy status can be identified in satellite images of the region (L.I. Wilschut et al. 2013; Wilschut et al. 2015). Potentially this would allow for the generation of a public health tool using only satellite images and no field records to anticipate periods of high *Y. pestis* prevalence. Technologies similar to this have been predicted to be increasingly utilised in future public health (Meerman 2008). The potential association between a burrow's connectivity and the likelihood that it will become occupied when empty could then be utilised to further improve these models, providing guidance as to where gerbils may be likely to move to over the coming seasons. As demonstrated here, parameterisation of such a model would require a substantial number of burrow's occupancy time series to be known, although a potential source for this could be the satellite images themselves rather than further field studies. Indeed, the size of the data required may be prohibitive of effective field data being generated.

Whilst there is likely an issue of power, the gerbil population size is known to fluctuate with cycles occurring over a number of years (Naumov & Lobachev 1975; Begon et al. 2006). The data collection period ranged over 2010-2013, during which time the gerbil population was at a low point following a crash (Davis et al. 2007b). The nature of the relationship between a burrow's connectivity and its subsequent use by gerbils may change depending on the current overall population. This may be harder to observe in periods of high gerbil numbers, where occupancy is more constant and there are failures to become occupied will be rare. Additionally, the time series generated for each burrow were not especially long, with only 6 season transitions in each representing around three breeding cycles/overwintering periods. In the sample here, it would be difficult to identify multiannual cycles.

In a separate study it was identified that environmental properties including the presence of sand and trees in the area immediately surrounding a burrow were associated with the burrow being, and remaining occupied (Levick et al. 2015). This is a contrast to the results described here, where no significant association between burrow specific properties or colonisation and extinction could be identified. A potential explanation may lie again in the population experiencing a current crash, that the burrows with consistent occupancy remain so throughout the period of lower

numbers, and so hetereogeneities in this are harder to observe and demonstrate statistically.

Beyond the presence of the gerbils themselves, an association between the connectivity of a burrow and colonisation events may then have implications for the movement of *Y. pestis*. For a well-connected burrow to become colonised suggests that there is frequent gerbil movement to it. So, not only through the increased likelihood of the presence of the animals themselves, but further any fleas that they may introduce, a well-connected burrow would be anticipated to be at a higher risk of hosting *Y. pestis*. This remains a topic for further study for a number of reasons, primarily because little to no *Y. pestis* infections were observed in this area at the time of the study. Further, gerbil movements have yet to be explicitly tracked, although again a new study beginning in Kazakhstan may begin to answer these questions (Yeszhanov, Pers. Comms.).

The degree of clustering of the burrows identified previously (Wilschut et al. 2015) led to the suggestion that a metapopulation of the gerbil population might consider clusters of burrows, rather than individual burrows as the patch. A spatial regularity was identified in the placement of burrows at a local scale, which was suggested to arise due to the territorial behaviour of the gerbils (Randall et al. 2005). However, Wilschut et al did also suggest that these relationships should be explored over a larger spatial scale to identify whether they still held. Here we consider the role of spatial structure of the gerbil population at a larger scale, and over multiple seasons. The above study could be repeated including differing spatial scales of individual units, and this may feed in further to studies considering the potential appearance of hotspots driving persistence. If gerbil movement behaviour is limited to a certain set of burrows (which may be true, given the previous study on geographical barriers (Liesbeth I Wilschut et al. 2013)), it may in fact comprise of a series of sub populations. Additionally, information on visitation events was not considered in this study. The manner in which this is integrated may need to be considered carefully, given the relatively higher uncertainty of the data, but this may be able to provide a more specific measure of gerbil movements.

Given that the percolation process identified to drive *Y. pestis* transmission through the population is in itself dependent on burrow connectivity (Davis et al.

2008), if such a similar relationship exists in the gerbil population it may be that the percolation action of *Y. pestis* is an artefact of percolation of gerbils through burrows. Or perhaps, a separate percolation effect over which the movement of *Y. pestis* acts as a further phenomenon. Demonstrating this would require both additional empirical and theoretical evidence, but this potential serves to highlight the role of spatial structure and environmental properties in the transmission of infection through this population. It may be that the connectivity relationship identified here may be capable of acting as a proxy for the previously described occupancy threshold, which may be better suited to the satellite imagery techniques considered earlier.

A novel aspect of the above work is the use of statistical model projection to generate new response data with a larger sample size, but relationships reflective of those in the collected field data. This is discussed further elsewhere (Levick, Begon and Cornell, in prep). In any field-based study there is often a question of small sample size obscuring potential relationships, due to innumerable barriers to data collection and the (often impassable) barrier of extending data sets once collection is finished. Analysis of these generated data sets allows for some putative exploration of whether potential associations have been obscured because of lack of power. This is of course a conclusion to be taken with great caution, but towards conducting further investigation (as is planned, here) do give some supporting evidence.

Previous studies have demonstrated the gerbil population of the pre Balkhash to have strong spatial and social structure to their population (Wilschut et al. 2015). Their population dynamics have been explicitly linked to the presence and properties of *Y. pestis* outbreaks (Davis et al. 2004; Davis et al. 2008; Reijniers et al. 2012). Here we find that the population dynamics cannot be accurately or even nearly replicated by a spatial patch occupancy model approach, but neither are they simply being driven by connectivity. However, simulations suggest that a larger sample size may clarify the latter relationship. We suggest this is evidence for further empirical study of the population dynamics, towards an improved understanding of how the gerbil population drives *Y. pestis* infection.

3.6. References

Alexander, H.M. et al., 2012. Metapopulations and metacommunities: Combining spatial and temporal perspectives in plant ecology. *Journal of Ecology*, 100(1), pp.88–103.

Begon, M. et al., 2006. Epizootiologic parameters for plague in Kazakhstan. *Emerging infectious diseases*, 12(2), pp.268–73.

Bolker, B. & Grenfell, B., 1995. Space, persistence and dynamics of measles epidemics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 348(1325), pp.309–20.

Davis, S. et al., 2007a. Empirical assessment of a threshold model for sylvatic plague. *Journal of the Royal Society, Interface / the Royal Society*, 4(15), pp.649–57.

Davis, S. et al., 2007b. Plague metapopulation dynamics in a natural reservoir: the burrow system as the unit of study. *Epidemiology and infection*, 135(5), pp.740–8.

Davis, S. et al., 2004. Predictive thresholds for plague in Kazakhstan. *Science (New York, N.Y.)*, 304(5671), pp.736–8.

Davis, S. et al., 2008. The abundance threshold for plague as a critical percolation phenomenon. *Nature*, 454(7204), pp.634–7.

Fournier, D. et al., 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. In *Optimum methods software*. pp. 233–249.

Hanski, I., 1994. A practical model of metapopulation dynamics. *Journal of Animal Ecology*, 63(1), pp.151–162.

Hanski, I., 1991. Single-Species Metapopulation Dynamics: Concepts, Models and Observations. *Biological Journal of the Linnean Society*, 42(1–2), pp.17–38.

Hanski, I. & Gilpin, M., 1991. Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of the Linnean Society1*, 42, pp.3–16.

Hanski, I. & Ovaskainen, O., 2003. Metapopulation theory for fragmented landscapes. *Theoretical Population Biology*, 64(1), pp.119–127.

Hughes, N. et al., 2015. What are the effects of density, demography and identity on host movements? Prelimary insights from recent work. In *"Advances in the prediction of plague outbreaks in Central Asia" Closing Meeting*. Almaty.

Jesse, M. & Heesterbeek, H., 2011. Divide and conquer? Persistence of infectious agents in spatial metapopulations of hosts. *Journal of Theoretical Biology*, 275(1), pp.12–20.

Keeling, M.J. & Gilligan, C.A., 2000. Bubonic plague: a metapopulation model of a zoonosis. *Proc Biol Sci*, 267(1458), pp.2219–2230.

Laperrière, V., Brugger, K. & Rubel, F., 2016. Cross-scale modeling of a vector-borne disease, from the individual to the metapopulation: The seasonal dynamics of sylvatic plague in Kazakhstan. *Ecological Modelling*, 342, pp.34–48.

Levick, B. et al., 2015. The perfect burrow, but for what? Identifying local habitat conditions promoting the presence of the host and vector species in the Kazakh plague system. *PLoS ONE*, 10(9), p.e0136962.

Levick B., Begon M., Cornell S., (in preparation). Simulations to explore the power of ecological field studies.

Levins, R., 1969. Some demographic and genetic consequences of environmental hetereogeneity for biological control. *Bulletin of the Entomological Society of America*, 15(3), pp.237–240.

Meerman, M., 2008. Space technology: a new frontier for public health. *Bulletin of the World Health Organisation*, 86(October 2005), pp.87–88.

Moilanen, A. & Nieminen, M., 2002. Simple connectivity measure in spatial ecology. *Ecology*, 83(4), pp.1131–1145.

Naumov, N.P. & Lobachev, V.S., 1975. Ecology of desert rodents of the USSR (Jerboas & Gerbils). In I. Prakash & P. K. Ghosh, eds. *Rodents in Desert Environments*. The Hague: The Hague, pp. 465–598.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. Available at: http://www.r-project.org/.

Randall, J.A. et al., 2005. Flexible social structure of a desert rodent, Rhombomys opimus: philopatry, kinship, and ecological constraints. *Behavioral Ecology*, 16(6), pp.961–973.

Reijniers, J. et al., 2012. A curve of thresholds governs plague epizootics in Central Asia. *Ecology letters*, 15(6), pp.554–60. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22449078 [Accessed October 31, 2012].

Reijniers, J. et al., 2014. Plague epizootic cycles in Central Asia. *Biology letters*, 10(6), pp.0–3.

Schmid, B. V. et al., 2012. Local persistence and extinction of plague in a metapopulation of great gerbil burrows, Kazakhstan. *Epidemics*, 4(4), pp.211–218.

Skaug, H. et al., 2013. Gneeralised Linear Mixed Models using AD Model Builder. R Package Version.

Venables, W.N. & Ripley, B.D., 2002. *Modern applied statistics with S.* Fourth., New York: Springer.

Wilschut, L.I. et al., 2013. Mapping the Distribution of the main host for plague in a complex lanscape in Kazakhstan : An object-based approach using SPOT-5 XS , Landsat 7 ETM + , SRTM and multiple Random Forests. *International Journal of Applied Earth Observation and Geoinformation*, 23, pp.81–94.

Wilschut, L.I. et al., 2013. Potential corridors and barriers for plague spread in Central Asia. *International journal of health geographics*, 12(1), p.49.

Wilschut, L.I. et al., 2015. Spatial distribution patterns of plague hosts: point pattern analysis of the burrows of great gerbils in Kazakhstan. *Journal of Biogeography*, 42(7), pp.1281–1292.

4. Distribution and dynamics of *Y. pestis* isolates lacking the fraction 1 (F1) antigen in a wild rodent population.

4.1 Introduction

The fraction 1 (F1) antigen covers the surface of the *Y. pestis* bacterium forming an extracellular pseudocapsule. Much is known about its expression and structure (Andrews et al. 1996; Du et al. 2002). The antigen appears to work in a synergistic manner with the pH 6 antigen (PsaA), also covering the surface of the bacterium, to inhibit bacterial uptake by host cells (Liu et al. 2006). In particular, the F1 antigen has been proposed as a virulence determinant, but contrasting results are observed (Burrows et al 1957; Donavan et al. 1961; Drozdov et al 1995; Samoilova et al. 1996; Weening 2011). However, the body of evidence concerning the immunogenicity of the antigen is clear, with a great number of studies demonstrating the F1 antigen is capable of generating protective immunity (Baker et al. 1952; Simpson et al. 1990; Andrews et al. 1996; Titball et al. 1997; Derbise et al. 2015; Galen et al. 2015)A full review of the related literature is provided in the appendices (Appendix A1).

In Kazakhstan, *Y. pestis* is found in wild great gerbil (*R. opimus)* populations in the desert and semi desert areas around Lake Balkhash in the south-east of the country. The F1 antigen is of particular interest to the study of the Kazakh system, as diagnosis of plague in the gerbils is currently conducted by detecting seropositivity to the antigen. *Y. pestis* strains that lack the F1 capsule (although the *caf1* operon is present) have been recovered from gerbils (Bektemirov et al. 1986). If the animals infected with $F1^-$ strains are present in sufficiently large numbers or in a clustered distribution, then they may significantly impact our ability to monitor plague dynamics in the region through generation of false negatives (demonstrated from a laboratory perspective by Li et al. (2008)). Beyond this body of work, evidence of frequent, or more sustained outbreaks of *Y. pestis* without the capsular antigen could also have implications for efforts to vaccinate humans at risk of contact with wild reservoirs, since F1/LcrV vaccines represent the majority of modern developments (Smiley 2008).

Whilst the genetics, structure and molecular role of the F1 antigen have been well characterised, few if any formal studies have attempted to describe the ecological or evolutionary impacts of the $F1^-$ phenotype at the level of the host population. It is

not yet known whether the F1$^-$ phenotype emerges or increases in frequency as a response to a particular change in the host or vector population. Moreover, it is not understood whether changes in the frequency of the phenotype lead to changes in the ecological dynamics of disease systems.

One notable complication in identifying evolutionary relationships is the large number of different genetic changes that may give rise to altered F1 production. However, as we are only able to consider observed phenotype, it is impossible to detect specific genetic or epigenetic processes occurring. Therefore here the ecological drivers and impacts of the phenotype are investigated but any evolutionary basis for these can only be considered hypothetically.

4.2 Aims and hypotheses

Aim

- To describe the distribution of and investigate the heterogeneity of the distribution of atypical F1- isolates of *Y. pestis* in a population of *R. opimus*

Hypotheses

- F1- isolates do not only appear as isolated, rare events.
- The distribution of the appearance of F1- isolates may be aggregated within space.
- The distribution of the appearance of F1- isolates may be aggregated in time.
- The distribution of F1- isolates is predictable, and this predictability could use to correct for the impact of such isolates on the ability to predict *Y. pestis* infection.

4.3 Methods

Field data collection and laboratory methods

The data described here were collected by Kazakh public health teams from 2000-2013 in the desert region south of Lake Balkhash in the South-Eastern of Kazakhstan. Field data include details of specific animals trapped: both *R. opimus* individuals (for whom life stage and sex were recorded) and insect vectors for individual burrows.

Alongside normal recording of *Y. pestis* presence across the Pre-Balkhash region, records were also generated detailing the F1 phenotype of isolates that were

recovered from the infected hosts and ectoparasites (fleas and ticks). *Y. pestis* strains were isolated from infected rodent hosts by plating blood and tissue samples on Hottinger's agar or from ectoparasites by crushing them in saline and plating on the same agar. *Y. pestis* identity was confirmed in the laboratory by examining morphology, growing speed, sugar fermentation and urea metabolism. Samples were confirmed as *Y. pseudotuberculosis* negative using the bacteriophage L-413C (Fillipov et al. 2012). Strains' ability to produce the F1 pseudocapsule was determined by double diffusion in gel, passive haemagglutination (PHA), and enzyme linked immune absorbent assay (ELISA) detecting antibody response (Martinevsky 1973; Bektemirov et al. 1986).

Records estimating overall gerbil occupancy levels were also collected by the public health teams in a separate study. These are recorded as the number of gerbils observed per sector (10 × 10 km) area, observed once in the spring season (May-August) and once in the autumn season (September-November) of each year.

Field accounts from natural plague foci in Uzbekistan and China have identified correlations between local levels of trace elements including iron, copper, cobalt and zinc in the environment and the location of epizootics in the foci (Gage & Kosoy 2005; Liu 2000; Mezentsev et al. 2000a). Similar observations have been made in plague positive areas of Tanzania (Meliyo et al. 2015). Additionally, laboratory studies have identified differential progression of *Y. pestis* infection in great gerbils subjected to diets with differential micronutrient profiles (Mezentsev et al. 2000b). With a view to identifying whether similar relationships exist in the Pre-Balkhash plague focus, perhaps related to the F1 phenotype, concentrations of six trace elements (zinc, cobalt, copper, iron, calcium, and manganese) were recorded at individual burrow locations across 6 study sectors in 2011 and 2012. These elements are known to regulate the expression of several *Y. pestis* genes (including virulence genes) (Carniel et al. 1987; Hu et al. 1998; Bearden & Perry 1999), and have been shown to determine host response to infection and infection outcomes across a range of mammal species, including rodents (Kubena & McMurray 1996). Specifically, the trace element composition at *R. opimus* burrow sites was recorded using plant stomach contents of the hosts, as these are representative of the true food intake, combining the hosts' food plants and any mixed soil consumed by the hosts. The details of the methods employed are detailed in the attached manuscript (Appendix A2).

Statistical Methods

The following analyses of the field and laboratory data were conducted in order to identify associations between the occurrence of the $F1^-$ *Y. pestis* isolates and potential drivers or predictors. The aim was to ascertain whether these heterogeneities were systematic, and therefore potentially could be related to a potential causative mechanism. When considering overall spatial or temporal trends, infection observations are considered as either all *Y. pestis* isolates or *Y. pestis* $F1^-$ isolates (rather than comparing $F1^+$ to $F1^-$ strains). This step was taken to identify where $F1^-$ strains behaved in a manner different to that observed by simply observing *Y. pestis* overall.

Data were available describing the distribution of all *Y. pestis* isolates (including $F1^+$ and $F1^-$) and atypical ($F1^-$) strains across the Pre-Balkhash plague system, for the period 2000-2013. As the records of *Y. pestis* isolates were relatively rare in these years, *Y. pestis* infection in gerbils, $F1^-$ strain isolation events, and gerbil occupancy were analysed at the level of the whole Pre-Balkhash area. Gerbil occupancy each year was calculated as an average of occupancy in all sectors that year (Levick et. al, 2015). The distribution of strains across isolates from *R. opimus*, from fleas (on and off rodents) and from ticks (similarly totalled) was calculated. Associations between the source of isolates and the appearance and strain types of *Y. pestis* were tested for, using Chi Square tests of Independence. In records of isolates from *R. opimus* individuals, the distribution of strains isolated from hosts was calculated for the two sexes, for adult and young individuals, and for females at different reproductive stages. Again, associations between these groups and appearance of *Y. pestis* and the appearance of the $F1^-$ phenotype were tested for using Chi Square tests of Independence.

In order to assess any potential associations between burrow location, year and either the isolation of any of the observed *Y. pestis* strains, or $F1^-$ strains alone, binomial logistic regression models were constructed. In both models only data from 2005-2013 were used, to reduce the influence of the earlier years where $F1^-$ strains were not isolated on the $F1^-$ models, and allow better comparison between identified relationships in the two models. In each case the models were reduced to a minimal model on the basis of Akaike's Information Criterion (AIC) and simplicity of the model. Terms would be deleted if they were considered to improve the model by a

significant change in the AIC (decrease by at least 2) (Burnham & Anderson 2002). Terms would be considered for deletion if they improved the AIC by a non-significant amount but would improve the simplicity of the model. In each case, latitude, longitude, year, and interactions between each of latitude/longitude and year were used as predictors in the initial model.

Burrows with data points over at least two years were used to investigate the potential relationship between the appearance of $F1^-$ isolates and previous occupancy by *Y. pestis* infected *R. opimus*. Each event was characterised as either *Y. pestis* absent, $F1^+$ strain present or $F1^-$ strain present. Although in chronological order, these events were not at regular time intervals, since burrows were not targeted to be trapped on a yearly basis. A Chi Square test of Goodness of Fit was used to identify whether the distribution of strains preceding $F1^-$ strain appearance differed significantly from uniform between the three possibilities (absence of *Y. pestis*, $F1^+$ strains and $F1^-$ strains).

A general trace element profile for each sector was generated by calculating average concentrations of trace elements from all observations within the sector in each year of data recording. These were reported as mean deviations from the mean across all sectors that year for that trace element, to minimise unrelated yearly variations. Taking these years' observations as an estimation of the sector's general trace element profile in comparison to other sectors, this was compared to the respective sector's *Y. pestis* infection time series and F1 time series to identify potential relationships between trace element levels and either *Y. pestis* or $F1^-$ strains emergence. Kendall's rank tests of correlation were performed to identify potential associations between sector's *Y. pestis* or $F1^-$ records and trace element profile.

4.4 Results

Temporal distribution

Over the 2000-2013 period, plague was seen in the pre-Balkhash region from 2001-2012, with F1$^-$ strains observed from 2005-2012 (Figure 4.1). Where F1$^-$ strains were observed they made up a minimum of 37% of the plague isolates, with all isolates found in 2006 lacking the capsular antigen.

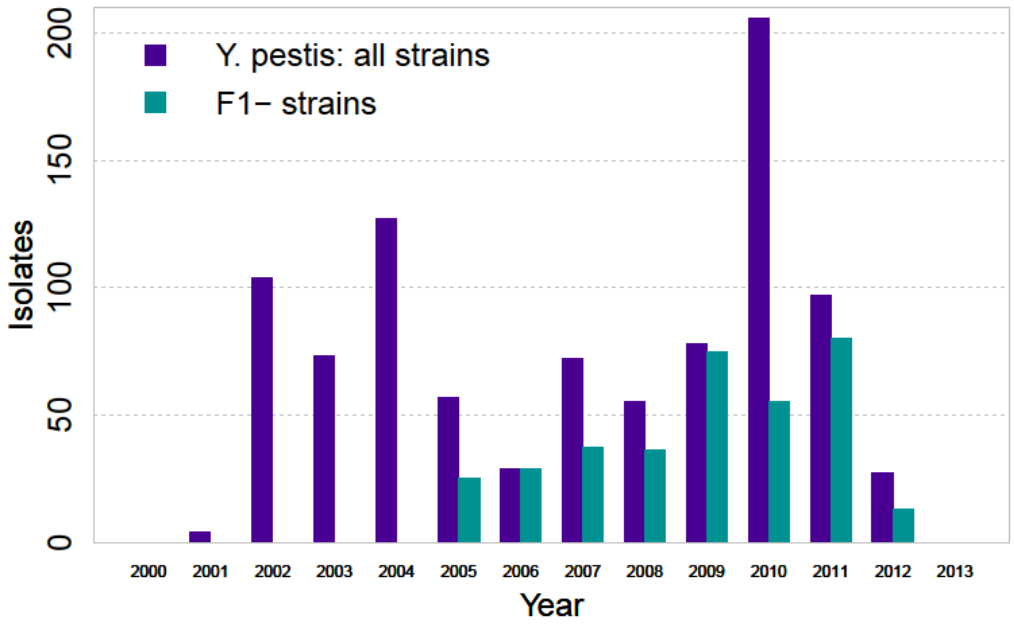


Figure 4.1. Numbers of all *Y. pestis* isolates, and F1- isolates, 2000-2013. No plague positive gerbils were recorded in these areas in 2000.

Spatial distribution

The spatial distribution of burrows with *Y. pestis* isolates and those where those isolates contained at least one F1$^-$ strain is shown in Figure 4.2. Both were observed across the whole trapping area.

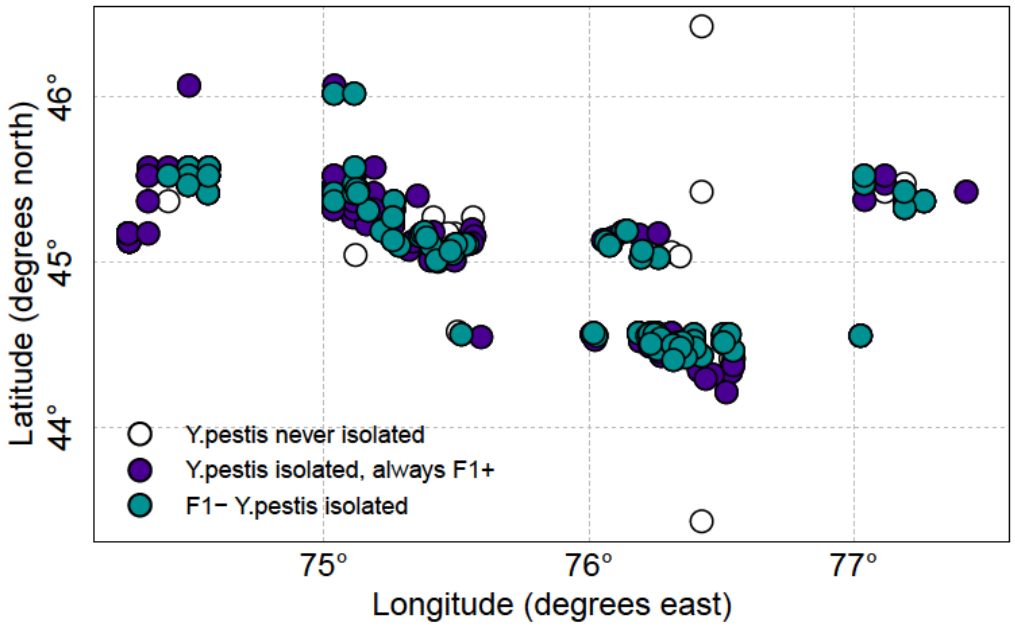


Figure 4.2. Locations of *R. opimus* burrows examined from records covering 1995-2013. Burrows where plague isolates were identified as having the F1- phenotype are shown in blue.

Binomial logistic regression models were constructed to assess potential relationships between location (as latitude/longitude position), year, and potential interactions between them with *Y. pestis* presence and $F1^-$ strain presence. For *Y. pestis*, the selected model showed weak, but significant associations between both latitude and longitude with *Y. pestis* presence. A slightly stronger, and significant, association was observed between year and *Y. pestis*, and the association between an interaction of year and longitude position was stronger again (although not reflected in an equivalent interaction with latitude, which was dropped from the model (Appendix A1)).

No significant associations were identified between the predictors and the presence of $F1^-$ strains (Table 4.1). The selection procedures are detailed in the appendix (Appendix A1).

Table 4.1. Minimal binomial logistic regression model output for the presence of *Y. pestis* of both phenotypes, and the presence of F1$^-$*Y. pestis*, 2005-2013. The model selection procedure is detailed further in the appendices (Appendix A1).

Model response	Factor	Coefficient Estimate	Standard Error	P Value
Y. pestis presence	Latitude (°N)	1.59	0.37	<0.001
	Longitude (°E)	-1684.06	237.02	<0.001
	Year	-63.81	8.99	<0.001
	Longitude:Year	0.84	0.12	<0.001
	(Intercept)	1.28×10^5	1.81×10^4	<0.001
F1$^-$ strain presence	Latitude (°N)	-74.93	219.77	0.73
	Longitude (°E)	-133.65	126.09	0.29
	Year	-6.70	8.90	0.45
	Latitude:Year	0.04	0.11	0.73
	Longitude:Year	0.07	0.06	0.29
	(Intercept)	13,500	17,800	0.45

Gerbil occupancy and infection with F1$^-$ strains

The time series of F1$^-$ isolates and gerbil occupancy appear to suggest that peaks in F1$^-$ occurrences follow peaks in gerbil occupancy (Figure 4.3). Hence a cross correlation analysis was performed using the occupancy time series, and the time series of the proportion of *Y. pestis* isolates that were identified as F1$^-$ strains (to ensure that any relationship observed was between F1$^-$ strains and occupancy, rather than the time series of *Y. pestis* occurrences overall). In the first four years of the time series no F1$^-$ strains were observed, and so comparisons were made only using the data from 2005-2013. There were no significant relationships (Appendix A1). There were also no significant associations with any of the different demographic groups of

gerbils (sex, age groups, or reproductive status in females; Chi-Squared test of Independence, all P values > 0.1) (Appendix A1). However, there was a significant association between the origin of the isolate and the strains identified (Table 4.2; Chi-squared test of Independence, df=6, X^2=40.55, P<0.001), with a greater proportion of $F1^-$ strains found in gerbils (and ticks) compared to fleas, although there was no difference between fleas caught on or off host (Chi-squared test of independence, df=2, X^2=1.70, P > 0.1).

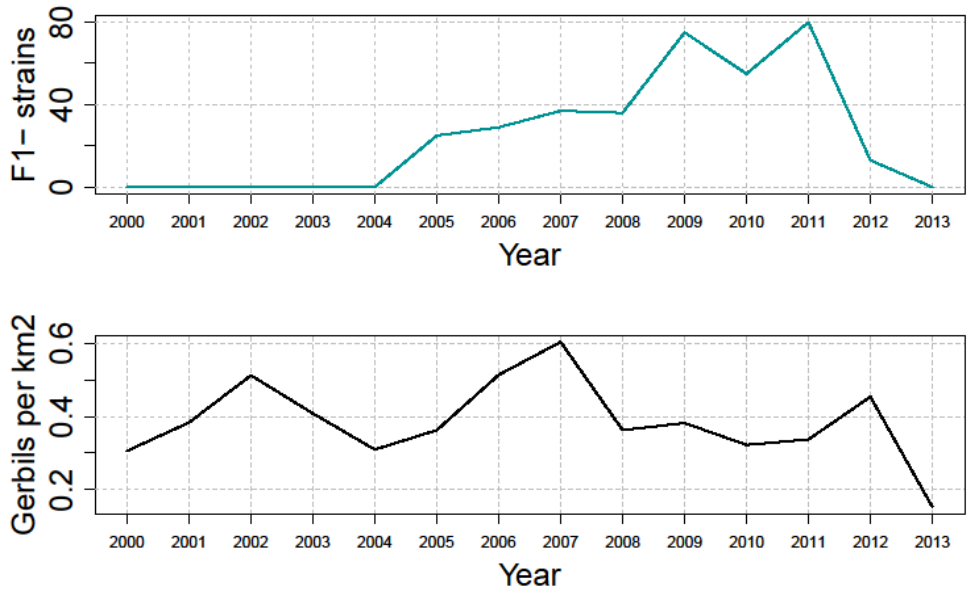


Fig 4.3. Number of rodents infected with an $F1^-$ strain and average gerbils recorded per square km.

Table 4.2. Distribution of $F1^+$ and $F1^-$ strains among different subsets of hosts. Chi Squared tests of Independence were performed 5, ignoring whether insects were migrating (off host) or on host animals.

Host isolate recovered from	***Y.pestis* not recovered**	**$F1^+$ strains**	**$F1^-$ strains**
Fleas (Migrating or on *R. opimus*)	110 (14.2%)	462 (60.0%)	204 (26.2%)
R. opimus	53 (18.6%)	117 (41.1%)	115 (40.4%)
Ticks (Migrating or on *R. opimus*)	12 (18.0%)	24 (35.8%)	31 (46.3%)

Representation of $F1^-$ strains amongst all Y. pestis isolate records

A similar analysis was conducted relating the number of *Y. pestis* isolates to the proportion of them with the $F1^-$ phenotype. Again the time series was reduced to 2005-2013, and again no association was identified at the level of the whole time series (Appendix A1). However, it appeared that $F1^-$ isolates were rarely seen occurring in a burrow without following the previous isolation of a $F1^+$ strain at the same burrow (Appendix A1). The frequency of these events and patterns of $F1^+$ and F1- strain appearance are given in Table 4.3.

In only 2 of the 36 occasions where $F1^-$ strains were isolated (5%) did their appearance follow absence of plague at the burrow, following both $F1^-$ and $F1^+$ strains with roughly equal frequency, (39% and 56% of $F1^-$ isolates respectively). This was found to be significant (Chi-Squared Goodness of Fit test, df=2, X^2=14, P<0.001). Moreover, in these two instances, when the distance between the burrow concerned and the nearest $F1^+$ *Y. pestis* occurrence in the preceding years was calculated, a burrow with $F1^+$ strains was identified within a close range. At one, a record of $F1^-$ strain occurred in 2007, following the last observation at the burrow in 2003 when no *Y. pestis* was recovered, but in the years 2003-2006 the nearest burrow with $F1^+$ was never more than 150 m away (see S2 File), well within ranges estimated for gerbil movement (Randall et al. 2005; Reijniers et al. 2012; Davis et al. 2008).

Table 4.3. *Y. pestis* isolation, and F1 expression phenotype of the isolates, following the last observation at a burrow being absence of *Y. pestis*, an $F1^+$ strain or an $F1^-$ strain.

***Y. pestis*/F1 status at first event in pair**	***Y. pestis* /F1 status at second event in pair**	**Number of observations**
Y. pestis absent	*Y. pestis* absent	1
	$F1^+$ strain	11
	$F1^-$ strain	2
$F1^+$ strain	*Y. pestis* absent	7
	$F1^+$ strain	27
	$F1^-$ strain	20
$F1^-$ strain	*Y. pestis* absent	2
	$F1^+$ strain	7
	$F1^-$ strain	14

Trace elements

Few burrows positive for *Y. pestis*, of either the $F1^+$ or $F1^-$ phenotype, were identified over the study period in the sectors where trace element data was collected (Table 4.4). Trapping sectors 7934 and 11742 have the largest number of observations to consider, and here we did observe higher levels of manganese than in the other sectors (Figure 4.4). There does appear to be a relationship between the ratios of iron:cobalt and iron:copper in each sector and their respective plague observations, particularly being higher in the south (sectors 10544 11742) where higher totals of plague observations are recorded (Figure 4.5). Kendall's rank coefficients were calculated for associations between each of the elements and both of the ratios

discussed above, with *Y. pestis* and F1$^-$ plague prevalence in the population. None were found to be significant (all P>0.1) (Appendix A1).

Table 4.4. *Y. pestis* and F1$^-$ isolates in all animals examined in the 6 sectors from 2000-2013.

Sector	**Number of burrows examined (2000-2013)**	**Number of burrows with *Y. pestis* isolates (2000-2013)**	**Proportion of burrow observations with *Y. pestis***	**Proportion of *Y. pestis* positive burrows with F1$^-$ type**
7934	20	19	0.95	0.68
9123	7	5	0.71	0.20
10512	2	2	1.00	1.00
10531	0	0		
10544	10	8	0.80	0.75
11742	27	12	0.44	0.17

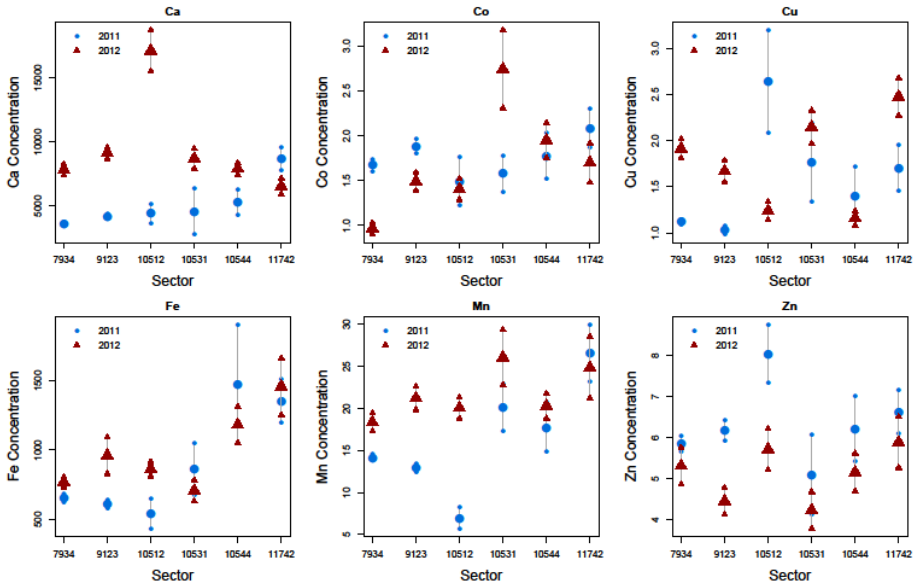


Figure 4.4. Concentration of trace elements in stomach samples from *R. opimus* individuals trapped across 6 sites (sectors). Results are presented as deviations from the mean values of that micronutrient across records from all sectors that year. Records from 2011 are shown in blue and 2012 in burgundy.

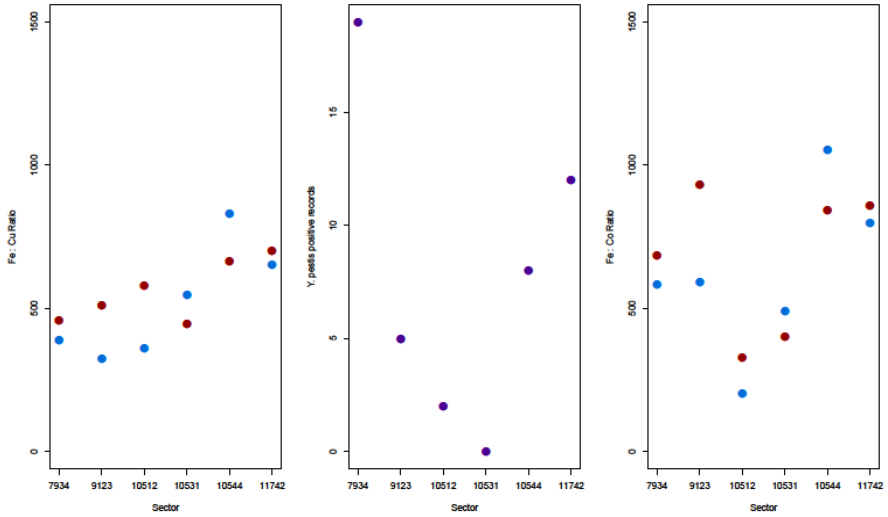


Figure 4.5. Concentration of trace elements in stomach samples from *R. opimus* individuals and *Y. pestis* prevalence across 6 sectors, 2011-12. Left to right: Iron to copper ratio in each sector, total *Y. pestis* positive animals observed in each sector, iron to cobalt ratio in each sector. Again, records from 2011 are shown in blue, and those from 2012 shown in burgundy.

4.5 Discussion

Atypical strains of *Y. pestis* lacking the F1 antigen have previously been isolated from human and other animal hosts (Winter et al. 1960; Williams et al. 1978; Anisimov et al. 2004). Viable strains lacking the capsular antigen can also be readily generated in the laboratory (Appendix A1). Whilst sometimes referred to as a "virulence factor" for plague, experimental studies give a mixed picture of the direct relevance of F1 to *Y. pestis* virulence (Anisimov et al. 2004), in some cases depending on the host genetic background (Weening et al. 2011). The high immunogenicity of the F1 capsule is reported repeatedly and has led to its utilisation as a vaccine component (Williamson et al. 1995; Heath et al. 1998), first tested in humans in the early 1970's (Meyer et al. 1974). Here several years of consistent identification of $F1^-$ isolates circulating in wild great gerbils, *R. opimus,* in the Pre-Balkhash desert of Kazakhstan is reported. From the current analysis, the pattern of occurrence of $F1^-$ strains appears to have no significant association with space, time or fluctuations in or subsets of the gerbil population. $F1^-$ isolates do however appear to effectively always arise in burrows with a history of previous *Y. pestis* presence. These results are presented as a first step towards a larger body of research of these atypical strains in the context of wildlife infections.

In the Ural-Emba region of Kazakhstan, $F1^-$ strains were isolated at the peak of an epizootic, representing 6.5% of isolates collected. This reduced to 0.64% in later periods of the epizootic. In other areas, isolates lacking the pFra plasmid have been reported to make up between 0.2-1.2% of the population of *Y. pestis* strains (Anisimov 2002). The likelihood of similar patterns being observed in other wildlife foci is supported by repeated observation of heterogeneous immune response profiles to plague within and between reservoir species (Andrews et al. 2010; Vernati et al. 2011). The absence of such observations is likely to have resulted from a lack of specifically directed effort to identify them in a population.

Differentiation of virulence phenotypes is a strategy observed in pathogenic organisms across all taxa, with clear alterations of the resulting host parasite relationships (Deitsch et al. 1997). Given the known ability of *Y. pestis* to remain in a focus through periods of low host numbers (Eisen & Gage 2009) we suggest that the $F1^-$ phenotype may arise through pressures generated when host numbers fall. The classical example of such a direct response in virulence to population change is that

following the introduction of the *Myxoma* virus into rabbit populations through the mid 20th Century (Edmonds et al. 1975; Best & Kerr 2000). However, from results of previous studies, the F1 capsule does not appear to have such a direct role in virulence (Donavan et al. 1961; Samoilova et al. 1996) (Appendix A1). Further, from our own results we are unable to identify any significant association between gerbil numbers or burrow occupancy and the proportion of the population with the $F1^-$ phenotype. It is of course possible that such a relationship is present but was not detected in the current study due to the low numbers or unsuitable distribution of data points such as repeat visits to infected burrows. Weak, significant relationships were identified between *Y. pestis* presence and geospatial location, year, and interactions thereof, which were not identified when looking at $F1^-$ strains in isolation. Whilst this may be indicative of differing spatiotemporal trends in $F1^-$ strain occurrence compared to all *Y. pestis* isolates, it is more likely here that the model is insufficiently informed, and the presence or lack of geospatial or temporal gradient related to $F1^-$ strains remains to be confirmed.

Alteration of surface antigens to counteract the adaptive immune system is a commonly recognised strategy in protozoa (Pasternak & Dzikowski 2009) and macroparasites (Philipp & Rumjaneck 1984). Here it was shown that effectively whenever an $F1^-$ strain was isolated in a burrow, it was preceded either by other $F1^-$ strains (the strain was established), or by $F1^+$ strains, but only very infrequently by absence of *Y. pestis* isolates. Without records of individual animal's immune status, it is not possible to know whether this is directly related to adaptive immune responses, but we propose that the loss of the F1 antigen may be a strategy for *Y. pestis* to reinfect previously infected - but immune - hosts, thereby effectively expanding the potential susceptible host population available to it. Records in the Russian literature indicate that $F1^-$ strains are able to infect hosts previously infected with typical *Y. pestis* (Anisimov et al. 2002). Here antibody titres were not recorded from animals. Identifying any correlation between these and infection with the atypical strains would represent a fundamental step towards understanding this relationship. Evidence from laboratory studies appears to suggest that *Y. pestis* isolates may be able to regain the $F1^+$ phenotype, suggesting that this could be a temporary strategy to mitigate the effect of changing selection pressures (Arsenyeva et al. 2014). Given the involvement of the F1 pseudocapsule in the development of biofilms in the flea vector, a further

complexity may arise here, as the loss of the capsule could have negative consequences for onward transmission due to a reduced tendency for 'blocking' in the flea (Sebbane et al. 2009). There was a significantly lower proportion of isolates from fleas having the $F1^-$ phenotype, but the population of bacteria in the fleas does not exist in isolation from those in the rodents and it is unlikely for traits such as this to be switched at each host entry event. We anticipate that to fully capture the dynamics of interactions at this scale would require individual level transmission studies.

A limitation of such an investigation to relate the presence of these atypical strains to either immunogenic, or virulence related effects is the lack of specific genotypic, and indeed phenotypic profiling. Classification of *Y. pestis* strains generally invokes a suite of markers, including the presence and size of the plasmids and the presence of key virulence associated genes (Perry & Fetherston 1997). To make any defintive statement about evolutionary processes associated with their presence would require further characterisation of the strains. Certainly, this is an obvious direction for future study. At present, it is not possible to revisit the strains collected here and conduct this analysis, as the facilities are not available in country, and the materials cannot be transported out of Kazakhstan.

Trace element profiles were identified to differ across *R. opimus* individuals trapped from different areas of the Pre-Balkhash region. Several micronutrients are essential for *Yersinia* function (Perry & Fetherston 1997) and previous studies in other foci have identified links between geographical trace element profiles and plague incidence (Liu 2000; Mezentsev et al. 2000a). Identifying and characterising these potential relationships between micronutrients and plague was identified as a central future direction for plague research (Gage & Kosoy 2005). Definitively describing the connection between the trace element data here with the *Y. pestis* profiles for the associated sectors must be done with caution, given the temporal difference and limited number. However, no statistically significant association between trace element concentration and either plague or $F1^-$ strain appearance could be established. Given the number and extent of our samples, and the relatively low presence of plague over this period, the host micronutrient access remains to be fully characterised. Indeed, outside of any association with F1- strains, the role of micronutrient content in the diet of *Y. pestis* hosts remains to be completely resolved across a number of wildlife foci. Other investigations would need to be performed to resolve questions

such as the host level interaction between these observations and their *Y. pestis* infection, such as circulating levels of micronutrients in the host blood (representing micronutrients absorbed), and markers of infection and disease.

The results presented above demonstrate that $F1^-$ *Y. pestis* has the potential to make up large percentages of plague epizootics observed in wild systems. Seropositivity to the F1 antigen is a commonly used test in host animals for plague status. It then follows that previous estimates of plague prevalence in this system (and probably others) are potentially underestimates, as animals only ever infected with F1- strains would appear as seronegative for plague under these tests. This may be further complicated by heterogeneity in host immune response by genotype (Li et al. 2008). Moreover, the degree of underestimation is itself likely to be variable, given the observation here that emergence of $F1^-$ strains tends to follow occurrence of $F1^+$ strains at the burrow level. Our findings suggest in particular that individual level data may be required to fully characterise the distribution of $F1^-$ strains present in populations and any effect they may have on epidemic dynamics. The above findings should therefore be regarded as a preliminary starting point for exploring these strains, and any impacts they may have on disease dynamics.

4.6 References

Andrews, G.P. et al., 1996. Fraction 1 capsular antigen (F1) purification from Yersinia pestis CO92 and from an Escherichia coli recombinant strain and efficacy against lethal plague challenge. *Infection and Immunity*, 64(6), pp.2180–2187.

Andrews, G.P. et al., 2010. Identification of In Vivo -Induced Conserved Sequences from *Yersinia pestis* During Experimental Plague Infection in the Rabbit. , 10(8).

Anisimov, A. et al., 2002. Molecular Genetics Microbiology and Virology - ResearchGate. *Molecular Genetics, Microbiology and Virology*, (4), pp.1–30.

Anisimov, A.P., Lindler, L.E. & Pier, G.B., 2004. Intraspecific Diversity of Yersinia pestis. *Clinical Microbiology Reviews*, 17(2), pp.434–464.

Arsenyeva, T. et al., 2014. Characteristics of plague agent strains not producing main capsular antigen F1 and approbation of separate methods of detection. (In Russian). *Universum*, (2), pp.81–87.

Baker, E.E. et al., 1952. Studies on Immunization Against Plague: I. The Isolation and Characterization of the Soluble Antigen of Pasteurella Pestis. *J Immunol*, 68, pp.131–145.

Batra, L. et al., 2014. HSP70 Domain II of Mycobacterium tuberculosis Modulates Immune Response and Protective Potential of F1 and LcrV Antigens of Yersinia pestis in a Mouse Model. *PLoS Neglected Tropical Diseases*, 8(12).

Bearden, S.W. & Perry, R.D., 1999. The Yfe system of Yersinia pestis transports iron and manganese and is required for full virulence of plague. *Molecular microbiology*, 32, pp.403–414.

Bektemirov, T., Zhouravleva, Y. & Litvinova, M., 1986. *Standards, strains, and metods of control of bacterial and viral preparations.* T. Bektemirov, Y. Zhouravleva, & M. Litvinova, eds., Moscow: Mechnikov Institute Press.

Best, S.M. & Kerr, P.J., 2000. Coevolution of host and virus: the pathogenesis of virulent and attenuated strains of myxoma virus in resistant and susceptible European rabbits. *Virology*, 267(1), pp.36–48.

Burnham, K.P. & Anderson, D.R., 2002. *Model Selection and Multimodel Inference* Second., New York: Springer-Verlag.

Burrows, T., 1957. Virulence of Pasteurella pestis. *Nature*, 179(4572), pp.1247–1247.

Carniel, E., Mazigh, D. & Mollaret, H.H., 1987. Expression of Iron-Regulated Proteins. Infection and Immunity, 55(1), pp.277–280.

Chanteau, S. et al., 2003. Development and testing of a rapid diagnostic test for bubonic and pneumonic plague. *The Lancet*, 361, pp.211–216.

Chanteau, S. et al., 1998. Plague , a Reemerging Disease in Madagascar. *Emerging Infectious Diseases*, 4(1), pp.101–104.

Cully, J.F. et al., 1997. Dynamics of plague in a Gunnison's prairie dog colony complex from New Mexico. *Journal of wildlife diseases*, 33(4), pp.706–719.

Davis, K. et al., 1996. Pathology of experimental pneumonic plague produced by fraction 1-positive and fraction 1-negative Yersinia pestis in African green monkeys (Cercopithecus aethiops). *Archives of pathology & laboratory medicine*, 120(2), pp.156–163.

Davis, S. et al., 2007. Plague metapopulation dynamics in a natural reservoir: the burrow system as the unit of study. *Epidemiology and infection*, 135(5), pp.740–8.

Davis, S. et al., 2004. Predictive thresholds for plague in Kazakhstan. *Science (New York, N.Y.)*, 304(5671), pp.736–8.

Davis, S. et al., 2008. The abundance threshold for plague as a critical percolation phenomenon. *Nature*, 454(7204), pp.634–7.

Deitsch, K.W., Moxon, E.R. & Wellems, T.E., 1997. Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiology and molecular biology reviews : MMBR*, 61(3), pp.281–293.

Derbise, A. et al., 2015. Complete Protection against Pneumonic and Bubonic Plague after a Single Oral Vaccination. *PLOS Neglected Tropical Diseases*, 9(10), p.e0004162.

Donavan, J.E. et al., 1961. Role of the capsule of pasteurella pestis in bubonic plague in the guinea pig. *Journal of Infectious Diseases*, Sep-Oct(109), pp.154–157.

Drozdov, I.G. et al., 1995. Virulent non-capsulate Yersinia pestis variants constructed by insertion mutagenesis. *Journal of Medical Microbiology*, 42(4), pp.264–268.

Du, Y., Rosqvist, R. & Forsberg, Å., 2002. Role of fraction 1 antigen of Yersinia pestis in inhibition of phagocytosis. *Infection and Immunity*, 70(3), pp.1453–1460.

Edmonds, J.W. et al., 1975. Myxomatosis: the virulence of field strains of myxoma virus in a population of wild rabbits (Oryctolagus cuniculus L.) with high resistance to myxomatosis. *The Journal of Hygiene*, 74(3), pp.417–8.

Eisen, R.J. & Gage, K.L., 2009. Adaptive strategies of yersinia pestis to persist during inter-epizootic and epizootic periods. *Veterinary Research*, 40, pp.1–14.

Fillipov, A. et al., 2012. Bacteriophages capable of lysing Yersinia pseudotuberculosis: efficiecy of plating tests and identification of receptors in Escherichia coli K-12. In A. M. P. de Almeida & N. C. Leal, eds. *Advances in Yersinia Research*. Springer, pp. 123–134.

Friedlander, A.M. et al., 1995. Relationship between virulence and immunity as revealed in recent studies of the F1 capsule of Yersinia pestis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 21 Suppl 2, pp.S178–S181.

Gabastou, J.-M. et al., 2000. An outbreak of plague including cases with probable pneumonic infection, Ecuador 1998. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 94(March 1998), pp.387–391.

Gage, K.L. & Kosoy, M.Y., 2005. NATURAL HISTORY OF PLAGUE: Perspectives from More than a Century of Research. *Annual Review of Entomology*, 50(1), pp.505–528.

Galen, J.E. et al., 2015. A Bivalent Typhoid Live Vector Vaccine Expressing both Chromosome- and Plasmid-Encoded Yersinia pestis Antigens Fully Protects against Murine Lethal Pulmonary Plague Infection. *Infection and Immunity*, 83(1), pp.161–172.

Heath, D.G. et al., 1998. Protection against experimental bubonic and pneumonic plague by a recombinant capsular F1-V antigen fusion protein vaccine. *Vaccine*, 16(11–12), pp.1131–1137.

Hu, P. et al., 1998. Structural Organization of Virulence-Associated Plasmids of Yersinia pestis Structural Organization of Virulence-Associated Plasmids of Yersinia pestis. *J Bacteriol*, 180(19), pp.5192–5202.

Kilonzo, B.S., Makundi, R.H. & Mbise, T.J., 1992. A decade of plague epidemiology and control in the Western Usambara mountains , north-east Tanzania. *Acta Tropica*, 50, pp.323–329.

Kubena, K.S. & McMurray, D.N., 1996. Nutrition and the immune system: A review of nutrient-nutrient interactions. *Journal of the American Dietetic Association*, 96(11), pp.1156–1164.

Levick, B. et al., 2015. The perfect burrow, but for what? Identifying local habitat conditions promoting the presence of the host and vector species in the Kazakh plague system. *PLoS ONE*, 10(9), p.e0136962.

Li, B. et al., 2008. Serologic survey of the sentinel animals for plague surveillance and

screening for complementary diagnostic markers to F1 antigen by protein microarray. *The American journal of tropical medicine and hygiene*, 79(5), pp.799–802.

Liu, F. et al., 2006. Effects of Psa and F1 on the adhesive and invasive interactions of Yersinia pestis with human respiratory tract epithelial cells. *Infection and Immunity*, 74(10), pp.5636–5644.

Liu, Y., 2000. *The atlas of plague and its environment in the People's Republic of China* Y. Liu, J. Tan, & E. Shan, eds., Beijing: Geographical Science of Natural Resources, Chinese Academy of Science.

Malkhazova, S.M. et al., 2014. Natural-focal diseases: mapping experience in Russia. *International journal of health geographics*, 13(1), p.21.

Martinevsky, I., 1973. Materials for typing natural foci of plague according to the genetic properties of the plague bacillus strains. *Jounrla of Hygiene, Epidemiology, Microbiology and Immunology*, 17, pp.272–278.

Meliyo, J. et al., 2015. Status and Variability of Soil Micronutrients with Landforms in the Plague Focus of Western Usambara Mountains, Tanzania. *International Journal of Plant & Soil Science*, 4(4), pp.389–403.

Meyer, K.F., Hightower, J.A. & Mccrumb, F.R., 1974. Plague Immunization. VI. Vaccination with the Fraction I Antigen of Yersinia Pestis. *The Journal of infectious diseases*, 129(May), pp.S41–S45.

Mezentsev, V. et al., 2000a. Experimental study of the influence of trace elements on plague infection. *Zh Mikobiol Moscow*, 1, pp.41–45.

Mezentsev, V. et al., 2000b. The effect of trace elements on the infecious process in plague in an experiment (In Russian). *Zhurnal Mikrobiologii*, 1, pp.41–45.

Parkhill, J. et al., 2001. Genome sequence of Yersinia pestis, the causative agent of plague. *Nature*, 413(6855), pp.523–527.

Pasternak, N.D. & Dzikowski, R., 2009. PfEMP1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite Plasmodium falciparum. *The international journal of biochemistry & cell biology*, 41(7), pp.1463–6.

Pavlovsky, E., 1946. Fundamentals of the theory of natural focality of transmissable human diseases. *Zh Obshsch Biol*, 7(1), pp.3–33.

Perry, R.D. & Fetherston, J.D., 1997. Yersinia pestis--etiologic agent of plague. *Clinical microbiology reviews*, 10(1), pp.35–66.

Philipp, M. & Rumjaneck, F.D., 1984. Antigenic and Dynamic properties of helminth surface structures. *Molecular and biochemical parasitology*, 10, pp.245–268.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. Available at: http://www.r-project.org/.

Randall, J.A. et al., 2005. Flexible social structure of a desert rodent, Rhombomys opimus: philopatry, kinship, and ecological constraints. *Behavioral Ecology*, 16(6), pp.961–973.

Reijniers, J. et al., 2012. A curve of thresholds governs plague epizootics in Central Asia. *Ecology letters*, 15(6), pp.554–60.

Riehm, J.M. et al., 2015. Diverse Genotypes of Yersinia pestis Caused Plague in Madagascar in 2007. *PLOS Neglected Tropical Diseases*, 9(6), p.e0003844. Available at: http://dx.plos.org/10.1371/journal.pntd.0003844.

Samoilova, S. V. et al., 1996. Virulence of pPst+ and pPst- strains of Yersinia pestis for guinea-pigs. *Journal of Medical Microbiology*, 45(6), pp.440–444.

Sebbane, F. et al., 2009. The Yersinia pestis caflMlAl fimbrial capsule operon promotes transmission by flea bite in a mouse model of bubonic plague. *Infection and Immunity*, 77(3), pp.1222–1229.

Simpson, W.J., Thomas, R.E.X.E. & Schwan, T.O.M., 1990. Recombinant capsular antigen (Fraction 1) from Yersinia pestis induces a protective antibody response in BALB/C Mice. *Americal Journal of Tropical Medicine and Hygiene*, 43(4), pp.389–396.

Smiley, S.T., 2008. Immune defense against pneumonic plague. *Immunological Reviews*, 225(1), pp.256–271.

Stenseth, N.C. et al., 2008. Plague: past, present, and future. *PLoS medicine*, 5(1), p.e3.

Titball, R.W. et al., 1997. Expression of the Yersinia pestis capsular antigen (F1 antigen) on the surface of an aroA mutant of Salmonella typhimurium induces high levels of protection against plague. *Infection and Immunity*, 65(5), pp.1926–1930.

Vernati, G. et al., 2011. Antigenic profiling of yersinia pestis infection in the Wyoming coyote (Canis latrans). *Journal of wildlife diseases*, 47(1), pp.21–29.

Weening, E.H. et al., 2011. The dependence of the Yersinia pestis capsule on pathogenesis is influenced by the mouse background. *Infection and Immunity*, 79(2), pp.644–652.

Williams, J.E. et al., 1978. Atypical plague bacilli isolated from rodents, fleas and man. *American Journal of Public Health*, 68(3), pp.262–264.

Williams, J.E., Moussa, M. & Cavanaugh, D., 1979. Experimental plague in the California Ground Squirrel. *Journal of Infectious Diseases*, 140(4), pp.618–621.

Williamson, E.D. et al., 1995. A new improved sub-unit vaccine for plague: the basis of protection. *FEMS immunology and medical microbiology*, 12(3–4), pp.223–230.

Wilschut, L.I. et al., 2013. Potential corridors and barriers for plague spread in Central Asia. *International journal of health geographics*, 12(1), p.49.

Winter, C.C., Cherry, W.B. & Moody, M.D., 1960. An unusual strain of Pasteurella pestis isolated from a fatal human case of plague. *Bulletin of the World Health Organization*, 23, pp.408–409.

5. Onchocerciasis associated epilepsy in the Democratic Republic of Congo

5.1 Introduction

The infectious agents classified as “Neglected Tropical Diseases” still cause a great deal of morbidity and mortality in human populations across lower income communities, especially across many countries in Sub-Saharan Africa (Hotez & Kamath 2009). Onchocerciasis is one such NTD, caused by infection with the parasitic helminth *Onchocerca volvulus*. *O. volvulus* is transmitted through bites of black flies of the family *Simuliidae*. In Sub-Saharan Africa, the flies are most often *Simulium damnosum s.l.*, but in the Democractic Republic of Congo (DRC) commonly *Simulium squamosum s.l.*(Krueger & Hennings 2006). Infection causes disease in humans with symptoms including itching and irritation of the skin, and damage to the cornea and sclera that ultimately may result in blindness, alongside a range of systemic symptoms. The pathology of *Onchocerca* infection arises from hyperimmune responses to the microfilariae (Simonsen 2009) released by adult worms that reside in the subcutaneous tissue after development from the injected L3 larval stage. These microfilariae are also the transmissible stage, picked up by *Simulium* when feeding on infected persons (Figure 5.1). More recently, it has been identified that the bacterial symbiont of *O. volvulus*, *Wolbachia pipientis* drives some of the pathological reactions when outside of the worm host (Brattig 2004).

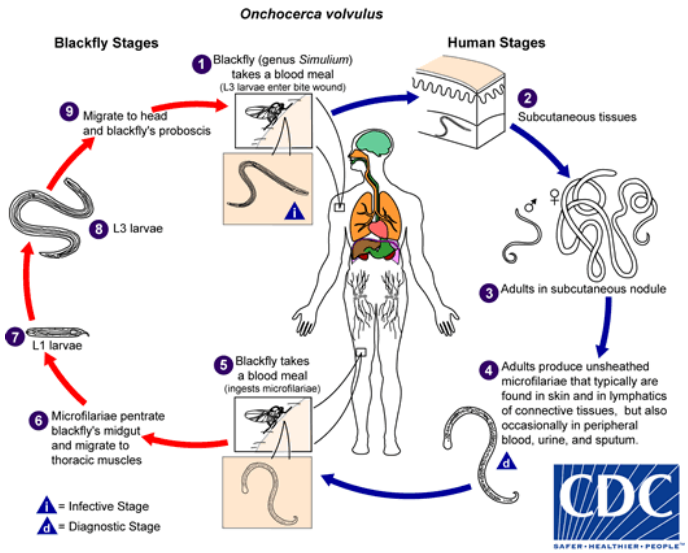


Figure 5.1. The life cycle of *Onchocerca volvlulus* (CDC 2013).

Ivermectin, a general antihelminthic, has been utilised for the control of onchocerciasis, paralysing and eventually killing the circulating microfilariae (Goa et al. 1991). This has the effect of reducing the associated symptoms, preventing the development of infection, and further helps to contribute to the prevention of onward transmission as the microfilariae are not available to be transmitted to the black fly (*Simulidae*) vector. However, despite the availability of the drug and two large international administration campaigns (first the West African Onchocerciasis Control Programme (OPC) (Boatin 2008) and later the African Programme for Onchocerciasis Control (APOC) (Coffeng et al. 2013)), many countries still have hyperendemic levels of onchocerciasis infection. These infections are concentrated in individuals living in rural areas surrounding the fast flowing rivers that allow for black fly breeding (Zouré et al. 2014).

In a number of areas where onchocerciasis remains hyperendemic in the population, the prevalence of epilepsy in the population has been observed to be atypically high (Pion et al. 2009). All these countries are located across central Africa, and fall into the Onchocerciasis Control Programme. The affected countries (Uganda, South Sudan, Tanzania and the Democratic Republic of Congo (DRC)) have also experienced recent and or ongoing conflict. Such disruptions have obvious repercussions on the health of the population in general, as well as the co-ordination and delivery of large public health programmes.

Moreover, an atypical form of epilepsy has been observed in these regions referred to as Nodding Syndrome (NS). NS is characterised clinically by atonic seizures of muscles in the head and neck, resulting in the head nodding forward repeatedly several times a minute (Dowell et al. 2013). Appearing in children aged 3-18 years, NS is often accompanied by hindered development, including stunted growth and failure to develop secondary sexual characteristics, although demographically those with NS do not differ to persons with epilepsy (Idro et al. 2014). The term “Onchocerciasis associated epilepsy” (OAE) is now often used to refer to a set of symptoms relating to epilepsy and NS observed in individuals in areas of high onchocerciasis endemicity (Colebunders et al. 2014).

This co-occurrence of onchocerciasis and epilepsy has been observed across a range of studies and countries (Kaiser et al. 2013). A recent meta-analysis of the

available literature estimated a 0.4% increase in epilepsy prevalence is predicted for a 10% increase in onchocerciasis prevalence (Pion et al. 2009). This relationship seems to exist across multiple countries, although it has not been shown definitively (Wamala et al. 2015). The exact functional relationship driving this remains uncharacterised.

Importantly, it has been demonstrated that other infectious agents known to cause epilepsy such as malaria (*Plasmodium spp.*) and cystercercosis (*Taenia solium*) alone are insufficient to explain the prevalence of epilepsy (Druet-Cabanac et al. 1999). Moreover, a case control study failed to identify significant difference in the prevalence of either between individuals with and without epilepsy (Colebunders, Mandro, et al. 2016a). Epilepsy is also associated with a great deal of social stigma, with children leaving school at a younger age, and impairments in movement and behaviour may lead to affected individuals coming to harm through accidents such as falling into rivers or fires, especially during seizures (WHO 2002).

That onchocerciasis is vector transmitted introduces additional layers of complexity in the epidemiology and ecology of both diseases, if it is the case that onchocerciasis is functionally related to the appearance of epilepsy. The fractured administration of Ivermectin across the regions of interest is also likely to play a pivotal, but potentially non-trivial role. Standard questions of epidemiology and risk factors also remain unanswered for OAE: what is the typical prevalence? How is this distributed demographically? A comprehensive measure of epilepsy morbidity, perhaps through estimating the loss of Disability Adjusted Life Years (DALY's) due to OAE also remains to be fully characterised, although this is beyond the scope of the present study.

The present study uses newly collected data from epileptic and non-epileptic individuals across the Orientale province in the north-eastern region of the DRC, to estimate the prevalence of epilepsy across the region, and to identify potential risk factors and epidemiological drivers. Overall it appears that significant challenges still remain in resolving the true epidemiology of OAE, but indications from the work so far lead towards localised drivers of differential prevalence. Both of these emphasise together the need for further, and more detailed investigation of the phenomenon.

5.2. Aim and hypotheses

Aim

- Describe the prevalence of epilepsy in a community in an *O. volvulus* endemic area
- Investigate ecological and population level factors that may indicate an association between onchocerciasis and epilepsy.

Hypotheses

- Epilepsy is present in the population, and at a higher prevalence than previously anticipated.
- Individuals with epilepsy are more likely to experience factors that increase their risk of being infected with *O. volvulus*.

5.3 Methods

Patient data was collected in two investigations: a case control study conducted in 2014 Titule, in the Bas Uele district of the Orientale province, and a larger prevalence study conducted over 2014-2015 over the Bas Uele, Tshopo and Ituri districts. The latter study is described here, and the former detailed elsewhere (Colebunders, Mandro, et al. 2016b). Additionally, the description of the methods concerning data collection (especially clinical details) are kept brief here, both for clarity and to make clear the present author's contribution to the work. Again this is detailed in the case control paper, and an earlier prevalence paper using an incomplete form of the data set discussed here (Colebunders, Tepage, et al. 2016).

Data collection

Over several visits between July 2014 and February 2016, 12,408 individuals were recruited from across the Orientale province (Figure 5.2). These belonged to 570 households, selected from villages with high levels of onchocerciasis endemicity based on nationally collected Rapid Epidemiological Mapping of Onchocerciasis (REMO) data (Noma et al. 2002).

Figure 5.2. Location of the health areas where the study sites are located, in the Orientalé province to the North East of the DRC.

Households were selected beginning at the centre of a locality, then approaching each third household. A household would only be used if eligible individuals were at home, and if not used the team would move to the next house. The households were geolocated using a handheld GPS device with ±4m accuracy.

In the Wanierukula health zone of the Tshopo province, a different sampling strategy was employed, which included all 1219 households along a 40km stretch of a national main road between the localities of Kisangani and Bunia. We refer to this region as the "PK30-PK70" as it does not exist, administratively.

Whichever sampling strategy houses were selected by, a one page questionnaire was conducted at each household detailing the age, sex and Ivermectin uptake in each year between 2000-2015 for every family member. Epilepsy was screened for using a five question questionnaire (Diagana et al. 2006). Family members would be asked to provide additional information if an individual was identified as potentially having epilepsy, about the circumstances and nature of the seizures experienced. Where a person was confirmed as having epilepsy, the age of first seizure was recorded (and month, if seizures appeared in the twelve months preceding the study date).

Additionally, extended interviews were conducted in a sub set of households to obtain data on the prevalence of skin lesions associated with onchocerciasis, referred to as “leopard skin”, indicating severe infections. In a smaller subset, the standard REMO procedure was conducted to determine onchocerciasis prevalence. Further records on onchocerciasis prevalence and regional level Ivermectin data was obtained from a national public health data base.

Statistical Analysis

Prevalence and risk factors

To investigate Ivermectin usage in cases compared to controls, a case control sample was generated from the survey data base using individuals with epilepsy who became so at the age of six years or above (meaning they were eligible for Ivermectin receipt at least one year before they became epileptic). Pairs were generated of individuals from the same health area, of the same age and gender. Matches were sought along a random vector of case individuals, and where multiple matches were identified the control partner was selected at random. Only cases where a full match could be identified were included in the sample. The association between epilepsy status and Ivermectin treatment history in this sample was explored through univariate binomial logistic regression models of Ivermectin receipt in the year immediately before epilepsy appearance (and the equivalent year in controls), and of the proportion of occasions in the years following epilepsy that Ivermectin was received.

To investigate this in the whole study sample, a generalised linear mixed effects model was constructed including fixed effect terms for the proportion of Ivermectin doses received in the years individuals were eligible and the presence or absence of onchocerciasis associated skin lesions. This was controlled with fixed effects terms for age and gender, and health area was included as a random effect to allow for differences in public health and risk factors at an administrative level.

A second generalised linear mixed effects model was constructed to identify associations between an individual receiving Ivermectin in a given year and location with their epilepsy status. This included an interaction term between year and location, to describe the location specific timeline of public health administration.

Spatial distribution

Occasionally in the data set, individual households were listed (with unique ID's) that shared a GPS location co-ordinate. This essentially indicates close neighbours, given the +- 4m accuracy of the GPS device used and information from the field on family distributions in households. Where this occurred, these points were grouped as one household for the purposes of the following analysis. Given these households typically belong to family groups, we consider this to be appropriate from an epidemiological perspective. Furthermore, in considering the true spatial position of the household, the error of the measurement indicates that the two households will be within overlapping 4m radii, so we can consider their likely location to be effectively the same.

Heterogeneities in the spatial distribution of households with individuals with epilepsy were assessed via the construction of intensity maps and estimation of clustering compared to randomly distributed points through K function estimation. Specifically, K function estimates were obtained for households with and without persons with epilepsy at 50m intervals from 200m to 1500m. At each interval, the difference between the K function estimate for the two groups of households was calculated and an envelope around these values estimated (Rowlingson and Diggle, 1993). This was performed for each of the main regions: Tshopo, Ituri and Bas Uele.

These summary analyses were also performed with the data from the PK30-70 region, considered as two distinct regions either side of the Babagulu forest reserve, referred to as the Eastern and Western forest regions (EFR and WFR, to the East and West of the reserve respectively)(Figure 5.3). This data was used to inform spatial models, as it had the best spatial resolution, with sampling of every household along a stretch of road.

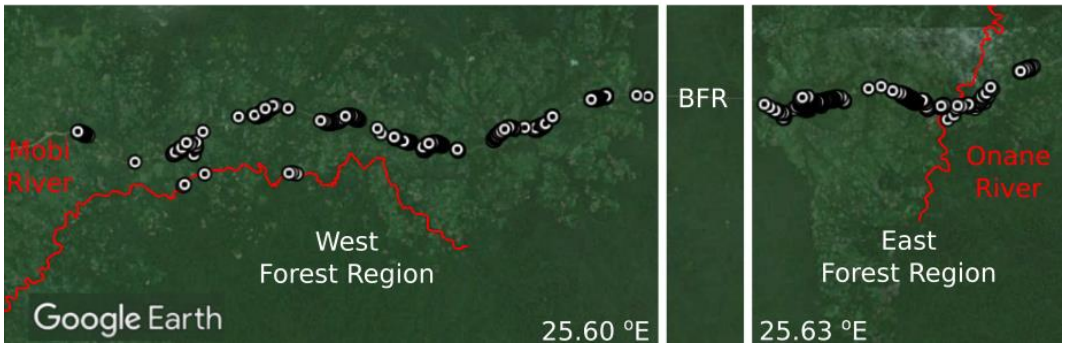


Figure 5.3. Households across the PK3070 region. The Babagulu Forest Reserve (BFR) lies between the Western Forest Region (WFR, up to 25.6°E) and the Eastern Forest Region (EFR, from 25.63°N).

To explore potential associations between risk factors (namely, the distance from the household to the nearest river, and the nearest forest edge), and the spatial distribution of the probability of a household having an individual with epilepsy binomial logistic regression models were constructed. A model was constructed with a smooth term describing the spatial location of the points (as geospatial co-ordinates). Several known associations were included to control the model: the proportion of males in the household, the average age of the household, and the Ivermectin administration in the region in the year 2014. A further model was constructed with the same predictive and control factors, but including a random effect describing administrative divisions, rather than a smooth term of location. The model with the smooth term allowed for identification of any smooth trend in spatial distribution underlying the prediction of locations of households with epileptics. Were this not identified, any spatial hetereogeneity could be sufficiently controlled for using the administrative region random effect term (Bivand et al 2013).

All analysis was conducted using the R statistical computing package (R Core Team 2014). Clustering and intensity analysis was conducted using “splancs” (Rowlingson & Diggle 2016). The mixed effects model with a smooth spatial term was fitted using “mgcv” (Wood & Scheipl 2013), and the mixed effects models with a random effect were fitted using “lme4” (Bates et al. 2015). Spatial distances were calculated using the R package “geosphere” (Hijmans 2016).

5.4 Results

Prevalence and risk factors

Of 12,408 people examined in the different health areas, 407 (3.3%) were found to have a history of epilepsy. Across the localities, households contained an average of six individuals (mean=5.65, s.d.=3.35). Whilst the median ages of the persons with epilepsy (18 years) and without epilepsy (16 years) were relatively similar, the age distribution of the persons with and without epilepsy was significantly different (Two Sample Kolmogorov-Smirnov test, D=0.26, P<0.001) (Figure 5.4).

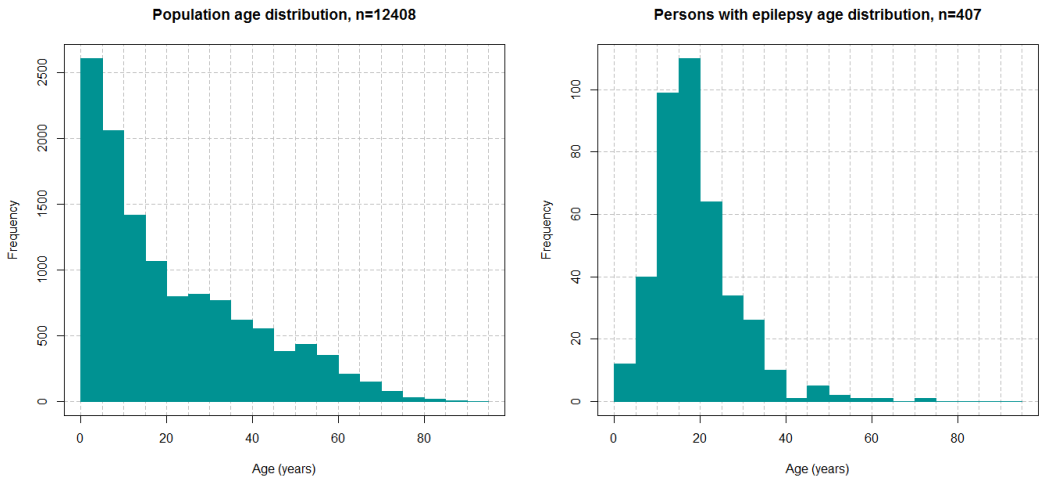


Figure 5.4. Histograms showing the distribution of ages in the whole population (left) and only those with epilepsy (right).

The median age of onset of epilepsy was 9 years, although the modal age of onset was 12 years. The highest prevalence of epilepsy was observed in the 10-19 years age group (Table 5.1).

Table 5.1. Age-specific prevalence rates of epilepsy.

Age (years)	Population of Household sample	N° of patients with epilepsy	Epilepsy prevalence (%)
> 40	2411	12	0.5%
30 – 39	1421	47	3.3%
20 – 29	1600	105	6.6%
10 – 19	2665	205	7.7%
< 10	4269	37	0.9%
All ages	12,366*	406*	3.3%

*42 individuals (of which one with epilepsy) had no age recorded and so are excluded.

Epilepsy appeared for the first time most frequently in the 10-15 year age group (Figure 5.5).

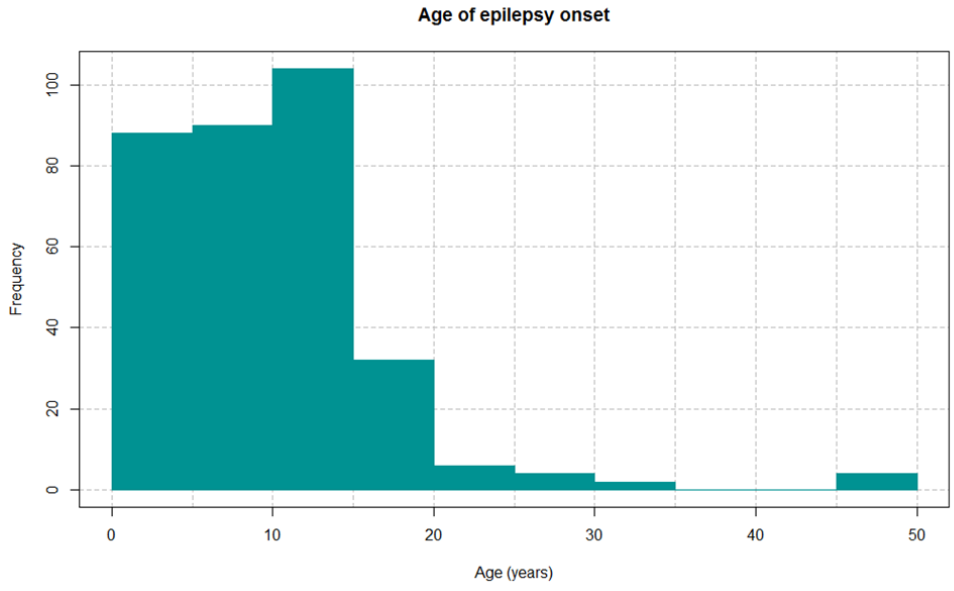


Figure 5.5. Histogram of age of epilepsy onset.

In 41 families there were at least 2 persons with epilepsy (1.5% of all households), and in 9 families at least three (0.3% of all households). Amongst the persons with epilepsy, 30% (123) lived in households with another person with epilepsy.

All persons with epilepsy had experienced seizures in the last 5 years. Nodding syndrome was not observed but persons with epilepsy with concomitant severe stunting and without external signs of sexual development were observed in the three provinces.

A high prevalence of epilepsy was observed in all health areas in the 3 provinces: ranging between 6.8-8.5% in Bas-Uélé, 0.8-7.4% in Tshopo and 3.6-6.2% in Ituri (Table 5.2).

Table 5.2. Prevalence rates of epilepsy, onchocerciasis (*O.v.*) endemicity, years of Ivermectin distribution and Ivermectin coverage in the surveyed health areas of Bas-Uélé, Tshopo and Ituri Provinces.

Province	**Health Zone**	**Health Area**	**Population Surveyed (Number of households)**	**Epilepsy prevalence**	**Epilepsy of recent onset[1]**	***O.v.* nodules[2]**	***O.v.* positive for OV16[6]**	**Years of IVM distribution**	**IVM coverage 2014**
Ituri	Logo	Draju	1063 (191)	6.2%	1.3%	22.2%[2]	0.5%	0	0%
	Rethy	Rassia	898 (200)	3.6%	0.3%	27.3%[2]	0.2%	3	87.8%
		Lokpa	846 (179)	3.7%	0.4%	13.8%[2*]	0.0%	3	73.0%
Tshopo	Mombongo	Yahuma	1259 (123)	2.1%	0.2%	57%[3*]	0.2%	10	33.0%
	Tshopo	Makutano Pumuzikia	203 (41)	7.4%	1.0%	94%[4*]	2.0%	9	17.6%
	Yaleko	Yatange	424 (60)	2.6%	0.2%	67%[4*]	4.0%	4	3.5%
	Wanierukula	Salambongo	5657 (524)	2.6%	0.2%	17.1%	1.9%	10	14.4%
	(PK30-70)	Makana	1121 (696)	0.8%	0	55%[4*]	0.4%	10	13.6%
Bas Uele	Aketi	Wela	570 (95)	6.8%	1.1%	68%[*]	9.7%	13	65.1%
		Makoko	367 (9)	8.4%	1.4%	68%[5*]	3.3%	13	78.1%

1) Onset of seizures in the last 12 months; 2) data obtained during our survey; *data obtained from the PLNO: 3) RAPLOA, 2005. 4) RAPLOA, 2003 and 5) RAPLOA, 2008 6) Positive test for OV16 antigens.

Between 2006 and 2014 on the other hand a larger proportion of epileptic individuals received Ivermectin compared with persons without epilepsy (Figure 5.6).

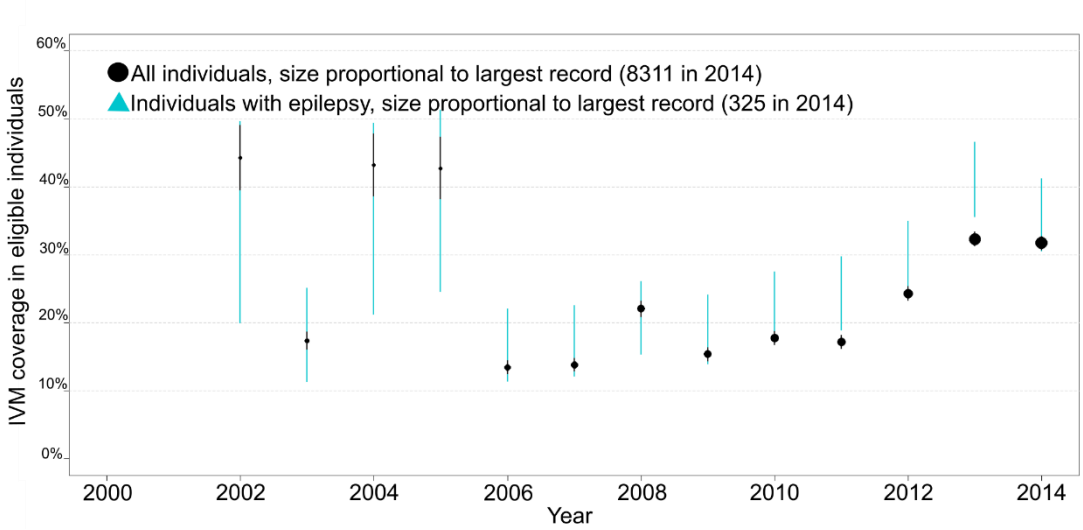


Figure 5.6. Proportion of epileptic and non-epileptic eligible individuals who were treated with Ivermectin, 2000-2014. The size of the point is proportional to the maximum number of individuals of the relevant epilepsy status dosed with Ivermectin in one year (the highest number of persons with and without epilepsy dosed in a given year was observed 2014). For the case control study, 114 pairs (228 individuals) were identified, split across the health areas as shown in Table 4.

For the case control study, 96 pairs (192 individuals) were identified, split across the health areas as shown in Table 5.3.

Table 5.3. Health zone and area distribution of case control sample.

Health Area	Number of case control pairs from this health area
Rassia	4
Lokpa	5
Mombongo	4
Salambongo	35
Makana	4
Wela	21
Makoko	14
Yatange	2
Makutano Pumuzikia	7
Total number of pairs	96

Within the case control pairs, ivermectin treatment history before the appearance of epilepsy in sufferers, compared to the same life period in unaffected controls, was found to be significantly lower (Table 5.4). Comparison of similar measures in the period after the appearance of epilepsy found no significant difference (P>0.05).

Table 5.4. Ivermectin history according to epilepsy status, in health area, age and gender matched case and control study univariate analysis, with a random effect for pair identity.

Factor	Odds Ratio	95% C.I.	P Value
Ivermectin received the year before epilepsy appeared (equivalent year in the controls)	0.52	0.28, 0.98	0.04
Proportion of occasions, in years eligible, Ivermectin received before epilepsy appearance (or equivalent period in controls)	0.47	0.22, 0.96	0.04

Using the whole survey data base, male gender, onchocerciasis skin lesions and being treated, at least once with Ivermectin, were associated with being a person with epilepsy (Table 5.5).

Table 5.5. Multivariate regression analysis of individual risk factors for epilepsy

Fixed effects		**Number in sample**	**Patients with epilepsy**	**Epilepsy prevalence**	**OR**	**95% CI**	**P Value**
Proportion of Ivermectin doses received					1.80	1.32, 2.28	0.02
Onchocerciasis associated skin lesions (leopard skin)	Present	205	18	8.78%	3.24	3.78, 2.70	<0.001
	Absent	12, 131	388	0.32%			
Age					0.99	0.98, 0.99	<0.001
Gender	Male	6041	221	3.66%	1.25	1.05, 1.45	0.03
	Female	6304	186	2.95%			
Health area random effect[1]					0.35		

1 Variance between random by-health area intercept

By contrast, considering the probability of an individual receiving Ivermectin (in each year when eligible) this was found to be negatively associated with being epileptic (GLMM, OR=0.83, 95% CI=0.01, 0.69, P<0.01). This was controlled for by fixed factors describing health area, year, and an interaction between the two, along with a random effect term for locality. The model was a poor fit, and these results should be taken with a great deal of caution.

Spatial distribution

Clustering

The position of households, and those with at least one person with epilepsy, were mapped across the three main regions (Figure 5.7).

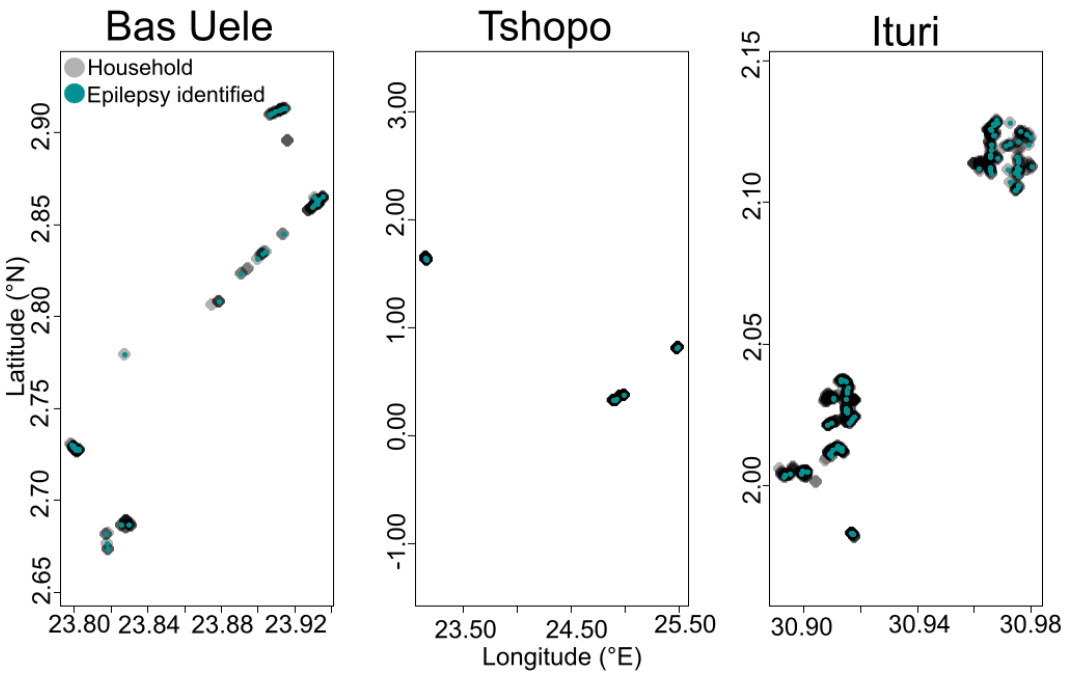


Figure 5.7. Geo positions of the included households in the three ecosites. Those in teal have at least 1 epileptic inhabitant.

K function analysis was conducted to assess clustering of households with persons with epilepsy compared to clustering of those with no sufferers. Some evidence of clustering was identified at very short distances in the EFR (<500m) (Figure 5.8).

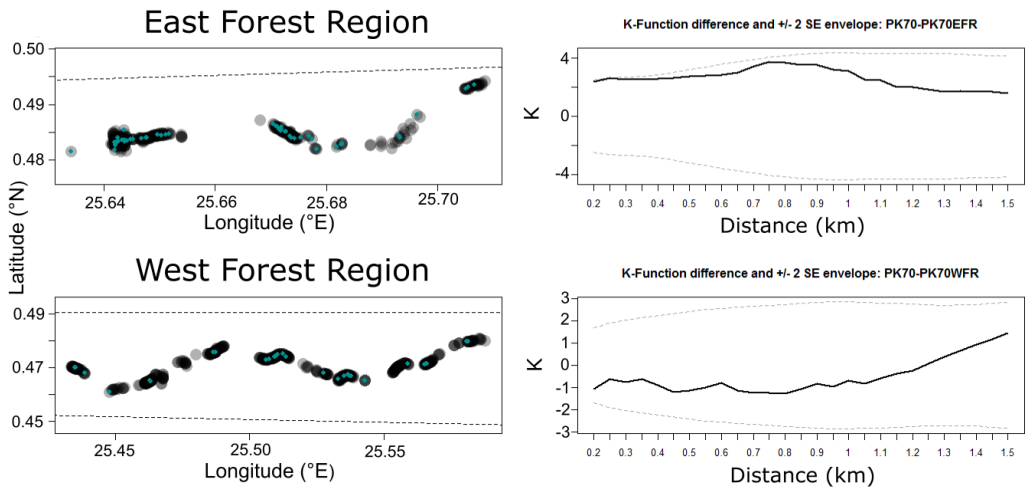


Figure 5.8. K function differences between households with and without epileptics, in the EFR (top) and WFR (bottom).

Clearer relationships could be observed by considering the intensity of households over the region with and without persons with epilepsy. In the EFR, a patch of high intensity of households with epileptics is observed in the centre of the region that is not reflected in the distribution of households overall, or the distribution of households without persons with epilepsy (Figure 5.9). This was not quite as distinct in the WFR (Figure 5.10)

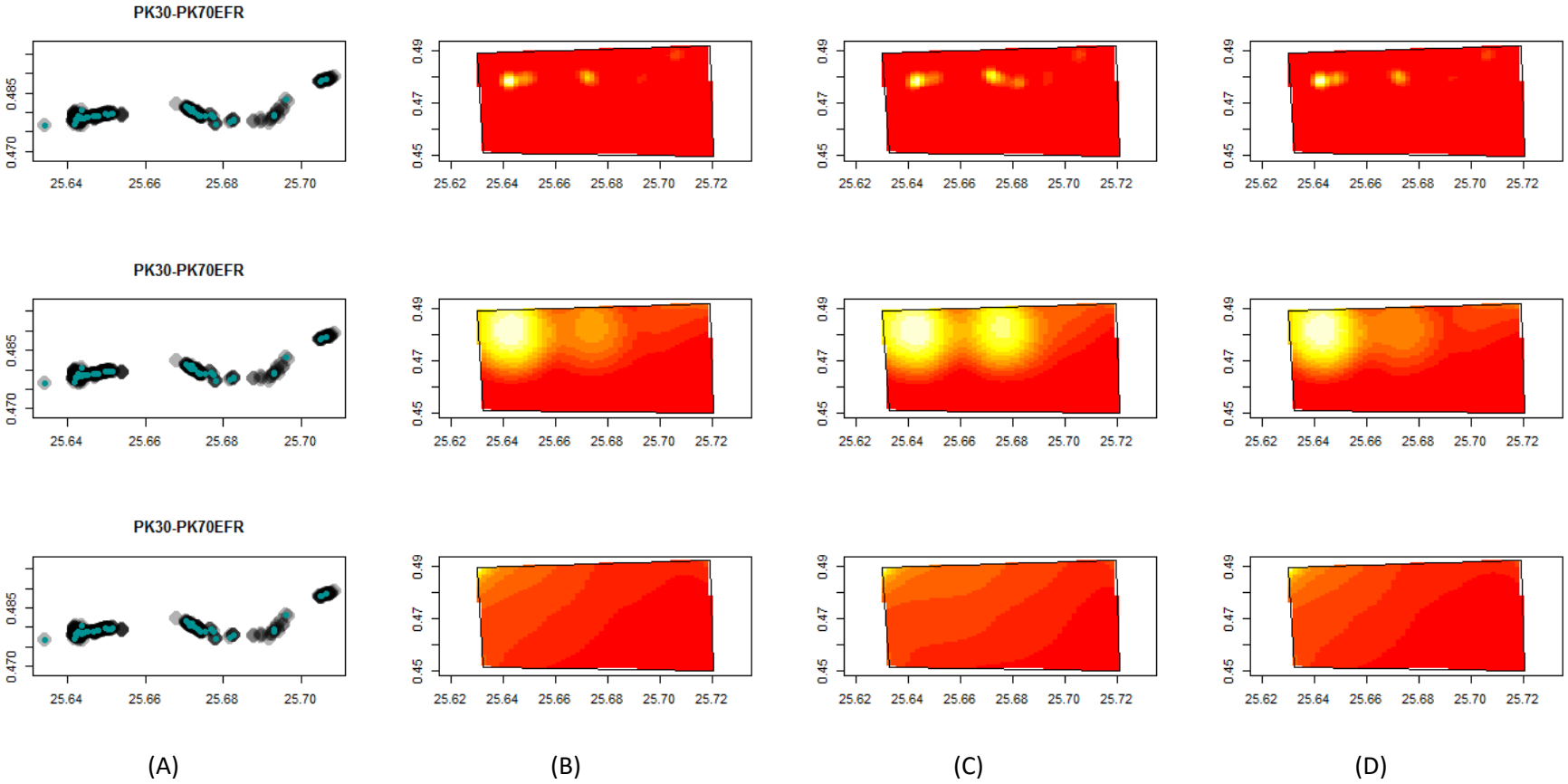


Figure 5.9. Intensity plot for the Eastern Forest Region, showing (A) a point map (as above), (B) intensity of households, (C) intensity of epilepsy positive households, and (D) intensity of epilepsy negative households. Each row shows the intensity plots at increasing bandwidths

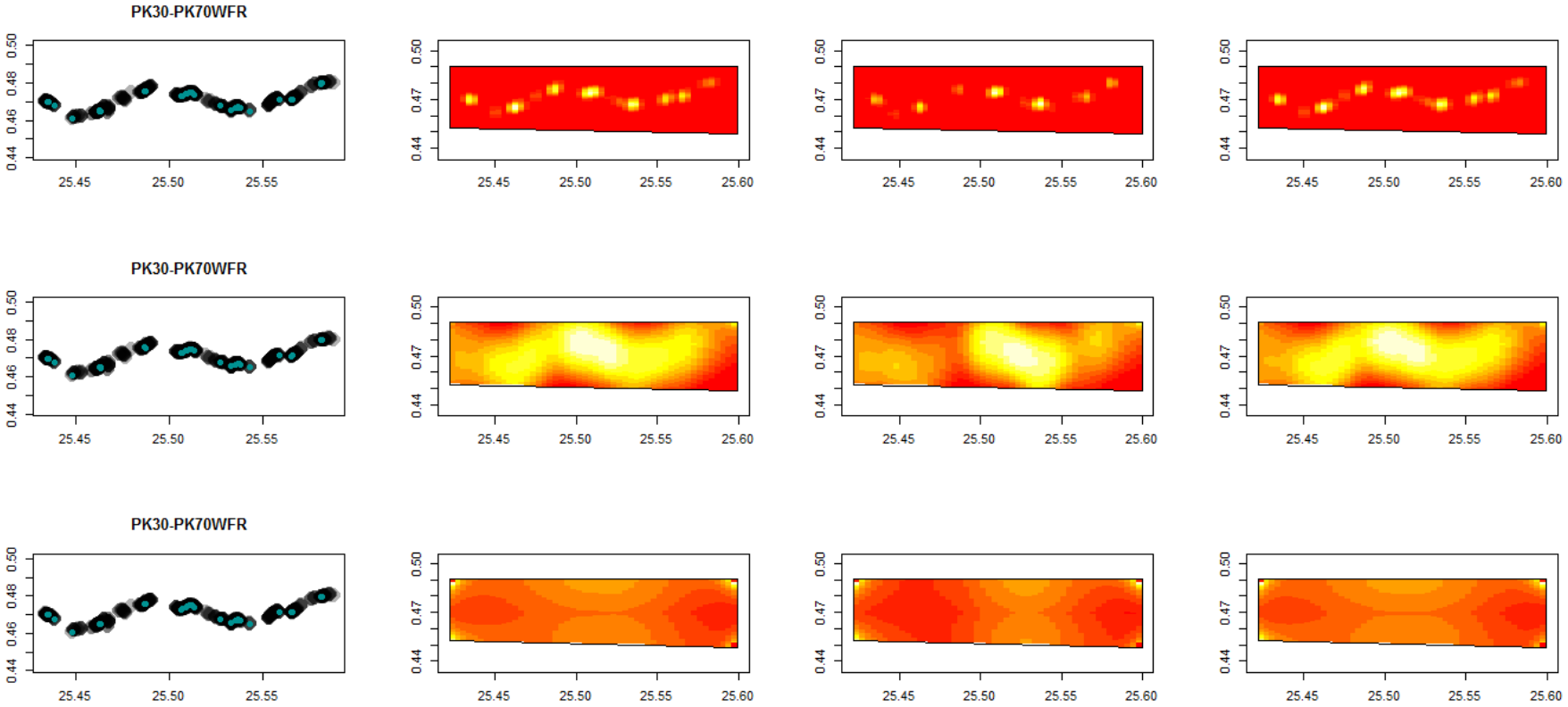


Figure 5.10. Intensity plot for the Western Forest Region, showing (A) a point map (as above), (B) intensity of households, (C) intensity of epilepsy positive households, and (D) intensity of epilepsy negative households. Each row shows the intensity plots at increasing bandwidths.

Binomial logistic regression models were then constructed to identify associations between a household having at least one person with epilepsy and putative ecological risk factors; distance to the nearest river, and to the nearest forest. The proportion in the household that are male, the average age of the household, and the ivermectin coverage in the locality to which a household belong were included as fixed factors. In the first instance, spatial trends were controlled for using a smooth term. This was found to be highly non-significant (P=0.23, Table 5.6). Given the lack of evidence for a smooth spatial trend, location was instead controlled for using a random effect term for the locality to which households belonged (Table 5.7). This identified a significant negative association between the distance from the household to the nearest river, and the probability of a household having a person with epilepsy.

Table 5.6. Generalised linear mixed model with a smooth term describing spatial location of points.

Factor	Estimate	Standard Error	P Value
(Intercept)	-3.08	2.57	0.23
Square root minimal distance to river	-0.51	0.79	0.52
Square root minimal distance to forest edge	0.17	0.37	0.64
Proportion of individuals in the household that are male	-0.03	0.58	0.96
The average age of the household	-0.01	0.01	0.28
Ivermectin coverage in the locality in 2014	8.58	5.92	0.15
Smooth spatial term			0.23

Table 5.7. Generalised linear mixed model with a random effect term for the locality to which households belong.

Factor	Estimate	Standard Error	P Value
(Intercept)	-0.39	1.13	0.73
Square root minimal distance to river	-0.68	0.28	0.02
Square root minimal distance to forest edge	-0.07	0.16	0.66
Proportion of individuals in the household that are male	-0.05	0.58	0.92
The average age of the household	-0.01	0.01	0.24
Ivermectin coverage in the locality in 2014	-2.76	4.17	0.51
Locality (random effect variance)	0.18		

5.5 Discussion

Unusually high epilepsy prevalence has been observed in countries with endemic levels of onchocerciasis infection, including Tanzania, South Sudan, Uganda and the Democratic Republic of Congo (Pion et al. 2009). Whilst associated with high onchocerciasis prevalence, and with some evidence of a functional link between the infection and epilepsy, the exact connection between the two remains to be resolved (Wamala et al. 2015; Kaiser et al. 2013).

The present study is an investigation of a large body of cross sectional data collected from across the Orientalé province of the DRC, with the intention of estimating prevalence of epilepsy in these areas, and identifying putative risk factors and characteristics of OAE specifically. A previous case control study conducted in the region showed significant differences in behaviour and symptoms between those with and without epilepsy (Colebunders, Mandro, et al. 2016b). This second, larger prevalence study, was unable to identify clear risk factors driving epilepsy on a large scale, but again reiterated the link between onchocerciasis and its wider ecology with epilepsy.

Ivermectin was identified to potentially mitigate risk, but in a manner dependent upon both location and birth year. Potential hotspots for epilepsy were observed, and spatial risk was identified to be unlikely to be best described as a smooth surface. Increasing distance from the nearest river appeared to have a significant, negative relationship with epilepsy risk in a household. Overall this demonstrates that the etiology of epilepsy prevalence is likely to be complex, and may involve different public health and ecological factors, possibly in an interactive manner. Ultimately it is likely that local relationships drive this unusual pattern of epilepsy, rather than large geographical gradients.

This study provides a further description of individuals affected by OAE. Evidence so far suggests that OAE and NS is typically found in younger individuals (Colebunders, Mandro, et al. 2016b; Winkler et al. 2014; Pion et al. 2009; Kaiser et al. 2013), and we see this pattern here again. Households with more than one individual with epilepsy make up around 1% of all households with epilepsy. For NS specifically, similar unpublished results, are discussed from a study in Uganda (Idro et al. 2016) and were reported previously from Tanzania (Winkler et al. 2008). This latter result may begin to point towards a potential genetic relationship, but one of few studies able to perform genomic analysis identified no epilepsy specific genes or consistent rare genes associated with epilepsy (Dowell et al. 2013). Further investigation could look to focus on these multiple epilepsy households, to identify if comparing these with households with one or no persons with epilepsy may demonstrate a gradient of risk factors. This could then point towards risk factors for epilepsy itself.

The case control study previously identified a clear risk between lowered Ivermectin use and increased epilepsy risk (Colebunders, Mandro, et al. 2016b). Considering analysis of our entire sample this was not reflected here. Rather, the reverse was true, and persons with epilepsy were identified to have significantly higher Ivermectin use. Similar results were obtained from a matched case control study in the Central African Republic (Druet-Cabanac et al. 1999). By contrast, when we consider only our matched case control pairs, a significant negative association is identified between Ivermectin use and epilepsy risk. It is important to note that the matching criteria create a very specific subsample from the larger body of data: for example, as they are matched to cases by birth year, which will create a generally

younger sample than all individuals in the study. Moreover, case control pairs are matched by both location and year of birth. This means that they will have experienced the same timeline of Ivermectin administration. Local effects such as missed areas or even missed individuals in a given area will have been experienced in the same manner by cases and controls and so only here are they fully controlled for. This is reiterated by the subsequent model concerning the probability of receiving Ivermectin given epilepsy status, although these results should be taken with caution as a poor fit was achieved due to lack of data and a relatively complex model. This then may suggest the importance of such local, fine scale relationships. Such relationships could be suggested simply by considering prevalence of epilepsy across the health areas as shown in Table 2. This showed hetereogeneity of prevalence, with no particular grouping by geography or geographical gradient. Considered together this does point towards local drivers of epilepsy risk. This would, arguably, reflect the epidemiological behaviour of onchocerciasis (Zouré et al. 2014), although it has not been conclusively linked to the same ecological drivers.

The suggestion that this may explain the discordant results in the whole sample is strengthened by the spatial results. These demonstrate a lack of consistent relationship between administrative areas, pointing to a heterogeneity across the study area. Looking to the PK30-70 region as an area of good quality, high resolution data, this is reflected at a local level: high intensity areas of people with epilepsy are observed that do not appear to just reflect the distribution of households overall. These could be potential hotspots for epilepsy. More conclusively, the term describing smooth spatial variation was far from significant in a model of risk of households having a person with epilepsy. This indicates that spatial variation may then arise at a larger unit scale than the individual household, pointing to perhaps gross ecological differences or implementation of different public health policy. A continuously varying trend would likely not give rise to such stark differences unless in a very fragmented and heterogeneous landscape. The previous case control study identified that people with epilepsy were discouraged from eating insects and goat meat as these were believed by the local population to contribute to the disease, demonstrating the clear social significance of epilepsy (Colebunders, Mandro, et al. 2016b; WHO 2002). Planned qualitative epidemiological studies may begin to explain this aspect further. That epilepsy might be explained by some sort of local geographical or social factor

has been identified previously (Druet-Cabanac et al. 1999), even over any association with onchocerciasis.

The spatial model does identify that risk of epilepsy decreases significantly with increasing distance from the nearest river edge. This does provide strong evidence for a functional association between onchocerciasis infection and epilepsy risk. Few studies have shown a "dose-response" style relationship between onchocerciasis and epilepsy (Pion et al. 2009; Kaiser et al. 2013), so severity of infection or repeated reinfection is not a likely sufficient explanation alone. A proposed next step for this data is to find households nearer to rapidly flowing sections of the river, and whether they are differentially likely to have inhabitants with epilepsy. Beyond definitive contact tracing events, this would give clearer evidence of this functional causative link.

Variant levels of prevalence and an imbalanced age distribution may suggest that the appearance of epilepsy would likely form part of atypical symptom, rather than be a definite outcome of infection. The previous study in this region found that individuals with epilepsy reported more frequent contact with the black fly vector, even when controlling for onchocerciasis infection (Colebunders, Mandro, et al. 2016b). Blackflies are competent vectors for other infectious agents including viruses, and cotransmission with filiarases may be aided by rupturing of the fly labrum by the worm (Cupp & Cupp 1997). One, or a combination of cotransmitted agents may worsen or alter onchocerciasis symptoms resulting in epilepsy. Further, the pathogenicity of onchocerciasis is driven by hyperimmune reactions both to the worm (Simonsen 2009) and to their symbiont *Wolbachia pipientis* (Brattig 2004). So far studies of OAE and NS have failed to identify evidence for microfilariae or other life stages crossing the blood brain barrier (as would be necessary to cause epilepsy), through failure to detect them in the cerebrospinal fluid (CSF) (Winkler et al. 2008). A more compelling argument for how onchocerciasis may drive neurological effects is through some worsening of the neuroinflammatory reactions. This could, indeed, be related to the black fly bite event: black fly saliva carries a variety of immunomodulatory factors (Gillespie et al. 2000; Cupp & Cupp 1997). But this is not the most likely explanation, given similar adverse reactions are not observed to other biting insects such as mosquitoes, sandflies and fleas, and moreover the disconnect between the short term release of these compounds and the chronic nature of epilepsy.

Some evidence for atypical immunity has been noted as the presence of increased leimodin 1 antibodies in the CSF of epileptic patients (Johnson et al. 2014). A study in Tanzania identified a higher frequency of antibodies against the human neural VGKC protein complex in patients with NS, although this was from a relatively limited patient sample (Dietmann et al. 2014). This hypothesis remains at an early stage, although a large study investigating the neuropathology of patients with NS in particular is going ahead in Tanzania over the next few years (Idro et al. 2016).

The high prevalence of epilepsy in onchocerciasis endemic regions of African remains to be definitively explained. Here we strengthen the relationship between the two diseases, given risk associations between Ivermectin use and epilepsy. This is not a clear picture however, and moreover we show that the relationship present is likely to operate at a local, rather than global scale. We can consider potential avenues for investigation in the future. A primary target should be the collection of qualitative epidemiological data, to investigate social factors that may act as drivers, and to guide better care and prevention of epilepsy. Collection of data of the type presented here should continue, as this is still very much an ongoing emerging public health problem. Investigations of the connection with onchocerciasis infection should continue to consider the whole disease system, to resolve the relationships between epilepsy risk and both ecological and clinical factors. A great deal remains then to understand of the phenomenon, but progress continues and there are potential leaps of understanding to be made in the near future.

5.6 References

Bates, D. et al., 2015. Fitting Linear Mixed Effects Models Using lme4. *Journal of statistical software*, 67(1), pp.1–48.

Bivand, R. et al., 2013. Applied Spatial Data Analysis with R. Second Edition. Springer.

Boatin, B., 2008. The Onchocerciasis Control Programme in West Africa (OCP). *Annals of tropical medicine and parasitology*, 102 Suppl(1), pp.13–7.

Brattig, N.W., 2004. Pathogenesis and host responses in human onchocerciasis: Impact of Onchocerca filariae and Wolbachia endobacteria. *Microbes and Infection*, 6(1), pp.113–128.

CDC, 2013. https://www.cdc.gov/dpdx/onchocerciasis/. *CDC DPDx*. Available at: https://www.cdc.gov/dpdx/onchocerciasis/ [Accessed December 27, 2016].

Coffeng, L.E. et al., 2013. African Programme for Onchocerciasis Control 1995-2015: Model-Estimated Health Impact and Cost. *PLoS Neglected Tropical Diseases*, 7(1).

Colebunders, R. et al., 2014. Nodding syndrome—a new hypothesis and new direction for research. *International Journal of Infectious Diseases*, 27, pp.74–77.

Colebunders, R., et al., 2016. Prevalence of River Epilepsy in the Orientale Province in the Democratic Republic of the Congo. *PLoS Neglected Tropical Diseases*, 10(5), pp.1–13.

Colebunders, R., et al., 2016a. Risk factors for epilepsy in Bas-UéléProvince, Democratic Republic of the Congo: A case-control study. *International Journal of Infectious Diseases*, 49, pp.1–8.

Cupp, E.W. & Cupp, M.S., 1997. Black Fly (Diptera: Simuliidae) Salivary Secretions: Importance in Vector Competence and Disease. *J. Med. Entomol*, 34(2), pp.87–94.

Diagana, M. et al., 2006. Dépistage de l'épilepsie en zones tropicales: validation d'un questionnaire en Mauritanie. *Bulletin de la Société de pathologie exotique*, 99(2), pp.103–7.

Dietmann, A. et al., 2014. Nodding syndrome in Tanzania may not be associated with circulating anti - NMDA - and anti - VGKC receptor antibodies or decreased pyridoxal phosphate serum levels - a pilot study. *African Health Sciences*, 14(2), pp.434–438.

Dowell, S.F. et al., 2013. Nodding syndrome. *Emerging Infectious Diseases*, 19(9), pp.1374–1384.

Druet-Cabanac, M. et al., 1999. Onchocerciasis and epilepsy: a matched case-control study in the Central African Republic. *American journal of epidemiology*, 149(6), pp.565–570.

Gillespie, R.D., Mbow, M.L. & Titus, R.G., 2000. The immunomodulatory factors of bloodfeeding arthropod saliva. *Parasite Immunology*, 22(7), pp.319–331.

Goa, K.L., McTavish, D. & Clissold, S.P., 1991. Ivermectin: A Review of its Antifilarial Activity, Pharmacokinetic Properties and Clinical Efficacy in Onchocerciasis. *Drugs*, 42(4), pp.640–658.

Hijmans, R., 2016. geosphere: Spherical Trigonometry.

Hotez, P.J. & Kamath, A., 2009. Neglected tropical diseases in sub-Saharan Africa: Review of their prevalence, distribution, and disease burden. *PLoS Neglected Tropical Diseases*, 3(8), pp.2–11.

Idro, R. et al., 2016. Is nodding syndrome an Onchocerca volvulus-induced neuroinflammatory disorder? Uganda's story of research in understanding the disease. *International Journal of Infectious Diseases*, 45, pp.112–117.

Idro, R. et al., 2014. Patients with nodding syndrome in Uganda improve with symptomatic treatment: a cross-sectional study. *BMJ open*, 4(11), p.e006476.

Johnson, T. et al., 2014. Detection of auto-antibodies to leiomodin-1 in patients with nodding syndrome. *Journal of Neuroimmunology*, 275(1–2), p.103.

Kaiser, C., Pion, S.D.S. & Boussinesq, M., 2013. Case-control Studies on the Relationship between Onchocerciasis and Epilepsy: Systematic Review and Meta-analysis. *PLoS Neglected Tropical Diseases*, 7(3).

Krueger, A. & Hennings, I.C., 2006. Molecular phylogenetics of blackflies of the Simulium damnosum complex and cytophylogenetic implications. *Molecular Phylogenetics and Evolution*, 39(1), pp.83–90.

Noma, M. et al., 2002. Rapid epidemiological mapping of onchocerciasis (REMO): its appliation by the African Programme or Onchocerciasis Control (APOC). *Annals of tropical medicine and parasitology*, 96(Supplement 1), pp.S29–S39.

Pion, S.D.S. et al., 2009. Epilepsy in Onchocerciasis Endemic Areas: Systematic Review and Meta-analysis of Population-Based Surveys. *PLOS Neglected Tropical Diseases*, 3(6), p.e461.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. Available at: http://www.r-project.org/.

Rowlingson, B. & Diggle, P., 2016. splancs: Spatial and Space-Time point pattern

analysis.

Simonsen, P.E., 2009. Filariases. In G. C. Cook & A. I. Zumla, eds. *Manson's Tropical Diseases*. pp. 1491–1500.

Wamala, J.F. et al., 2015. Nodding Syndrome May Be Only the Ears of the Hippo. *PLOS Neglected Tropical Diseases*, 9(8), p.e0003880.

WHO, 2002. EPILEPSY : A manual for Medical and Clinical Officers in Africa and Clinical Officers. *World Health Organisation*, p.124p.

Winkler, A.S. et al., 2014. A longitudinal study on nodding syndrome - A new African epilepsy disorder. *Epilepsia*, 55(1), pp.86–93.

Winkler, A.S. et al., 2008. The head nodding syndrome - Clinical classification and possible causes. *Epilepsia*, 49(12), pp.2008–2015.

Wood, S. & Scheipl, F., 2013. gamm4: Generalized additive mixed models using mgcv and lme4.

Zouré, H.G.M. et al., 2014. The geographic distribution of onchocerciasis in the 20 participating countries of the African Programme for Onchocerciasis Control : (2) pre-control endemicity levels and estimated number infected. *Parasites & vectors*, pp.1–15.

6. Discussion

Infectious agents, from viruses to multicellular parasites, are a marker of their host species' interactions with the ecosystem that it inhabits. These interactions may have been established in evolutionary past, with some microbes and hosts evolving to share symbiotic, even mutualistic interactions (Moran et al. 2008). Others retain an antagonistic relationship with their hosts, and new parasites and pathogens continue to emerge and move into new host species (Jones et al. 2008). Here two infectious agents of concern to human public health are considered from an ecological perspective, in order to inform our understanding and ultimately aid prevention and control of their associated diseases. For both new insights are generated, but equally many new questions and avenues for further research emerge. Together, they serve to highlight the role of ecology in infectious disease, and emphasise that a holistic approach to epidemiology and disease ecology can offer a fully rounded and more informed perspective of disease.

Yersinia pestis is considered first, the causative agent of the disease plague in humans. The bacterium is maintained in reservoirs of wild tolerant rodent species, from where it can spill over into populations in which it can cause harm. Here the dynamics of the host species and the distribution of atypical strains of the bacterium are examined. New properties of the population dynamics are identified and the atypical strains found to be widespread in both time and space.

Secondly, areas endemic for infection with *Onchocerca volvulus* in humans, recently reporting higher than expected prevalence of epilepsy in the population are investigated. The potential drivers for the epilepsy cases, specifically the link between onchocerciasis presence and epilepsy, are investigated. The present study is able to strengthen current evidence linking onchocerciasis with epilepsy risk, and identify that local patterns of behaviour, rather than global risks, are likely drivers of epilepsy.

Wildlife Yersinia pestis infections in Kazakhstan

Previous models of *Y. pestis* dynamics in Kazakhstan were unable to consider the real location of individual burrows or dynamics at this scale. Rather they rely upon random artificial networks of gerbil burrows that met empirical measures of real communities, such as gerbil density (Davis et al. 2004; Davis et al. 2007a; Reijniers et

al. 2014; Schmid et al. 2012; Laperrière et al. 2016). In Chapters two and three, the dynamics of the gerbil population at this resolution are considered for the first time, using field records of individual burrows collected between 2011 and 2013. In Chapter two, it is identified that environmental properties of burrows are good predictors of an individual burrow being, and staying occupied, and further of the presence of fleas at the burrow (Levick et al. 2015).

In particular, the presence of trees was found to be associated with a burrow being occupied, and the presence of sand with a burrow remaining occupied for longer periods. Both properties are likely to be related to the generation of a suitable microclimate for the burrow. Burrows constructed by *Meriones crassus* with more sand present in the sediment had a higher average temperature (Shenbrot et al. 2002) which may aid overwintering and the maintenance of long term occupancy. The presence of trees would suggest a minimum level of moisture in the surrounding area or the presence of shade, known to be favoured by other fossorial species (Naumov & Lobachev 1975; Kausrud et al. 2007).

Equally, conditions suitable for trees may be generated or aided by the gerbils themselves, by their digging activity creating a feedback of micronutrients resulting in a "fertile island" effect (Xu et al. 2012; Whitford & Kay 1999). Negative associations between occupancy and the presence of clay soils may relate to the ease of digging in such soils. Overall it is important to understand that, at the level of the individual burrow, environmental properties can have a significant relation to the presence of gerbils (and indeed, of fleas, although a primary driver for this was not clearly identified here). This suggests then, that in anticipating areas likely to harbour sufficient gerbils and fleas for *Y. pestis* to be present, these environmental properties should be considered.

Exactly how this is factored in to current projections is not resolved here. A potential approach is explored in Chapter three, but with limited success. It is identified that Spatial Patch Occupancy Models (SPOM's) (Hanski & Ovaskainen 2003), of a standard form at least, are not sufficient to project *R. opimus* dynamics. Including variant colonisation probabilities to reflect seasonality in the gerbil population dynamics was unable to resolve the discrepancy between the projections and observed data. It may be that the model is missing a key, perhaps unidentified,

property of gerbil population dynamics that is not reflected. However, a further investigation into the relationship between colonisation probability and the connectivity of burrows again was unable to find significant associations. Through the use of simulations, it was proposed that this may be due to insufficient sample size, both in terms of number of burrows and of seasons covered. The data available here provide occupancy time series for 513 burrows over 7 seasons, representing 6 season to season transitions. To properly inform models of population dynamics, it may be the case that data covering a larger number of seasons is required. In particular here only 3 over winter periods are available, likely important for determining population numbers. Furthermore, other than estimates of movement range, little is known about gerbil movement behaviours.

A new study is currently underway in Kazakhstan to investigate gerbil movement behaviour, amongst other things. The insights provided here may be able to inform analysis of this data towards resolving fine resolution dynamics of the gerbil population. Were this able to be fully characterised, this could be a powerful tool for informing statistical projections of outbreaks, perhaps in conjunction with information from satellite imagery (Wilschut et al. 2013). In light of the findings from Chapter two, considering the dynamics at this resolution is still likely to be useful, but this Chapter reveals how well informed such analysis would need to be.

In Chapter four, the phenotype of the *Y. pestis* isolates present in burrows are considered. The presence of atypical strains lacking the fraction 1 (F1) surface antigen at individual burrows was recorded and their distribution described. These strains were found to be common, both spatially and temporally, with no significant hetereogeneity in their distribution. Whilst no clear driver could be identified for the presence of these strains, this remains the first sylvatic foci in which the presence of F1- isolates has been considered from an ecological perspective, and one of few where they are discussed at all. Given the potential role for the F1 capsule in the virulence of *Y. pestis* (Drozdov et al. 1995; Davis et al. 1996; Weening et al. 2011) and its high immunogenicity (Andrews et al. 1996; Baker et al. 1952; Derbise et al. 2015; Galen et al. 2015; Simpson et al. 1990; Titball et al. 1997) the presence of these isolates, especially at as high a frequency as observed here, it would be surprising them for them to not effect infection outcomes and dynamics. It is of practical interest to understand their appearance as the presence of antibodies to the F1 antigen is typically

used to diagnose animals as being *Y. pestis* positive across many investigations of sylvatic foci. No clear conclusion could be drawn here for the presence or persistence of the atypical strains in the population. However, that the isolates were almost exclusively identified in burrows that had previously experienced typical *Y. pestis* infections suggests that there could be a link with reinfection of seropositive animals. Further questions certainly remain about these isolates. Fundamental information such as the genotypes present, and the fitness of these atypical isolates are not available at present. Equally, whilst the isolates are absent of the F1 antigen, it is not known if another structure is present in its place (that may serve some of the functions performed by the capsule). These questions remain for all sylvatic *Y. pestis* reservoirs, and an integrated lab and field approach would be required to fully describe the isolates themselves as well as their potential impact on the dynamics of *Y. pestis* in the host populations.

Onchocerciasis associated epilepsy in the Democratic Republic of Congo

Moving to human disease, in Chapter five potential drivers for the presence of onchocerciasis associated epilepsy (OAE) are investigated for a community in the Democratic Republic of the Congo (DRC). Much as was the case previously (Kaiser et al. 2013), the picture remains unclear with no definitive central aetiology. However, further understanding of the relationship between epilepsy risk and Ivermectin dosing leads to the conclusion that epilepsy epidemiology is likely dependent upon local, rather than large global relationships. The direction of association between Ivermectin dosing and epilepsy risk is effectively reversed when controlling for the local community to which an individual belongs, as well as their birth year. This demonstrates the importance of the specific time line of public health access to which individuals have been exposed.

Given the importance of reinfection in the establishment of symptoms and onward transmission of *Onchocerca volvulus*, these year by year relationships will be important to overall risk. Spatially, epilepsy risk appears in a manner better described as hotspots than a continual surface of risk over space. This is similar in many ways to onchocerciasis epidemiology (Zouré et al. 2014) but as yet the specific functional relationship is not known. Both *O. volvulus* itself, and the transmitting black fly vectors could be introducing biological artefacts resulting in atypical establishment of

symptoms, perhaps through divergence of existing hyperimmune reactions (Idro et al. 2016).

Synthesis

A broad set of conclusions (and just as many questions) are therefore reached from across the two distinct disease systems. However, both bring together interactions happening across the ecosystem to which they belong, and highlight the role of scale and resolution in approaches to understanding questions of disease transmission. This perspective is clearly seen when considering the system from an ecological perspective, demonstrating the utility of ecological field data collection and analysis in public health.

In the case of *Y. pestis*, the issue of scale arises from what is meaningful to consider as the unit of the individual. Previously it was identified that an individual animal's infection status does not need to be considered to make estimates of population level dynamics, given the speed and efficiency of within burrow transmission (Begon et al. 2006; Davis et al. 2007b). This work represented the first opportunity to inform predictions of gerbil occupancy over time (towards anticipating *Y. pestis* presence) with burrow specific information, rather than aggregate information such as gerbil or burrow density. Spatial analysis of occupancy records from single seasons revealed that burrows, and further occupied burrows, were clustered across the landscape (Wilschut et al. 2015). Wilschut *et. al* were, however, unable to resolve two considerations arising from this analysis: whether it would be similarly represented over a larger scale, and whether the clustering suggested a unit of individual consisting of several burrows.

Here contrasting results are identified, directed towards answering these hypotheses. In Chapter two, we see the importance of burrow specific environmental properties and combinations thereof for the presence of gerbils at burrows. However, in Chapter three, we were unable to generate projections of gerbil population dynamics that were reflective of our observed data through simulations of patch level dynamics or statistical models using the landscape connectivity of each burrow. As discussed previously, this may be an artefact of insufficient sampling to properly perform the analysis, especially in the number of seasons observed (Appendix A). On the other hand, a natural progression of this work that is unfortunately beyond the

current scope would be to repeat the analysis using individual study units comprising of multiple burrows, perhaps informed by spatial point pattern analysis as utilised by Wilschut *et al.*. It may be of interest to return to the data in Chapter two and concentrate on combinations of properties, rather than just their presence or absence. If the combinations do not tend to matter as much, simply that particular properties are present, this might suggest that what was observed here was in fact multi-burrow trends in environment type.

Our understanding of the presence of atypical *Y. pestis* isolates may also be further informed by such analysis. Are the isolates more common in particular environments, or is their appearance related to aggregations of burrows? In this instance however, many questions remain about the isolates' fitness, full phenotype and genotype, and within host dynamics. These questions remain not only for *Y. pestis* isolates in Kazakhstan, or indeed even just for wildlife isolates, but for any *Y. pestis* infection. Given that plague remains a public health concern (WHO 2013), there are broader reasons for understanding this aspect of the bacterium's biology.

In the case of epilepsy, rather than the unit of the individual we consider the spatial area of effect. High rates of epilepsy have been reported from across a number of countries, including the DRC (Pion et al. 2009; König et al. 2010). A number of studies have attempted to identify key risk factors, or indeed a cause of these unexplained cases of epilepsy. Whilst the correlation between onchocerciasis endemicity and high epilepsy prevalence has been noted on several occasions, but a single functional aetiology remains to be described (Druet-Cabanac et al. 1999; Newell et al. 1997; Colebunders et al. 2016). The more generalised term of "Onchocerciasis associated epilepsy" (OAE) has been used to describe the phenomena in less specific terms, and signifies the wider set of complications that could be associated with the observed patterns of epilepsy appearance (Wamala et al. 2015). Indeed, in the results presented in Chapter five, the conclusion begins to emerge that the explanation for OAE may not take the form of a single risk factor relationship. Given the additional considerations of a vector borne disease, and one where pathogenic effects are in large part generated through destructive hyperimmune reactions of the host themselves a wide breadth of biology must be considered in seeking to understand such links.

A continuous spatial risk surface for epilepsy in this study region could not be successfully resolved. Rather, especially in light of the quite stark differences in prevalence between different settlements, it appears more likely that epilepsy risk may be associated with locally varying drivers. We must consider the limitations of the present work: that the geospatial statistical model was constructed using only a subset of the data, and that sampling was not consistent (nor necessarily always related to population size) in the different localities studied. This second point is not ideal, but especially given recent political unrest in the DRC the realistic scope for data collection should be considered (and in this context, the recent effort is a significant one). This does perhaps point to more structured, standard epidemiological approaches to be taken in the future where possible, potentially including formal investigation of the role of Ivermectin.

It does also remain a possibility that the subset used in the models was unsuitable for identifying a continuous spatial trend. Two main arguments stand against this. The first is that some continuous trends were identified, in the relationship of epilepsy risk and distance from rivers, that existed outside of those related to spatial location. Secondly it is again reiterated that the prevalence differed so clearly between different localities where sampling took place over the whole data set. The identification of a significant association between epilepsy status and receiving Ivermectin (in years when eligible to do so) may suggest an explanation through different public health practises being observed in different administrative areas. Certainly anecdotally, researchers in the field report distinct heterogeneity of in-country medical staff opinions of the costs and benefits of Ivermectin (Laudisoit & Colebunders, pers. comms.). A next step may be to concentrate on identifying whether this difference in public health practise can be identified quantitatively from health records or through qualitative research. An alternative next step, that is planned for another group to take ahead, is to include more information on environmental differences such as the presence of rapids. At the very least, for the moment it unfortunately unlikely that a single, simple risk factor for epilepsy will be identified, and associated risk mitigation strategies implemented. There will be a great deal to learn from current, and planned laboratory work of human and *Simulidae* tissues.

Although approached from a different aim, both the investigations here serve to highlight the relevance of spatial scale and resolution in identifying patterns of

behaviour in infectious disease. Both studies were able to consider relationships occurring at different spatial and individual unit scales. In the *Y. pestis* study this leads to the question of the relevant scale of the individual in the system. This is a question of interest across infectious disease epidemiology, with approaches to anticipating disease ranging from standard epidemiological class (SIR, SEIR) models to increasingly population individual (agent) based models (IBM). A question of such models that remains without a universal solution is the appropriate measurement and representation of contact structures between individual units through which transmission can occur (Salathe & Jones 2010).

In the case of *Y. pestis* in great gerbils, it may be useful to take a step back from that question, to identify the level at which transmission and processes driving it are relevant to epidemiological outcomes. If it is the case that the potential future investigations discussed above identify that groups of burrows can actually be considered as the individual, then measurements and modelling of contact structure require less input. This would be an interesting conclusion in general for wildlife infectious disease, and perhaps raise the question of perhaps how habitat behaviour (here, burrowing) or environmental properties may affect epidemiological dynamics of infectious disease. This is of wider interest given the theoretical expectations of infectious agent behaviour in fractured populations (Hanski & Ovaskainen 2003), and moreover that this was failed to be observed in the Kazakhstan system (Jesse & Heesterbeek 2011).

Equally, that local, rather than global drivers appear to be driving the appearance of epilepsy may be a useful starting point for future investigations. This would not seem to fit with a universal relationship between any of the identified drivers and epilepsy: higher or lower than expected Ivermectin administration may not be sufficient to identify those at risk for epilepsy, neither would repeated *Onchocerca* infection. Investigation could then be guided to factors that may change at this administrative scale: the exact patterns of Ivermectin administration, ecological groupings of the black fly vector, or even human genetics. Indeed, in the case of the black fly-*Onchocerca* relationship, pairings of particular biotypes of each are known to have varying degrees of transmission success, and operate at a local scale (Basáñez et al. 2009). The case for human genetic association is weak (Dowell et al. 2013) but

may perhaps lie in some form of atypical immune response generating atypical pathologies from the *Onchocerca* infection (Johnson et al. 2014; Simonsen 2009).

The one health movement has revolutionised the approach to the research of human, animal and zoonotic diseases (Gibbs 2014). Ecology, and disease ecology is as integral a part of that discussion as veterinary, zoological, medical and anthropological expertise. The two investigations discussed here highlight the insights that can be offered of key epidemiological processes affecting infectious diseases by considering their ecological context. Where the prediction and control of infectious disease remain ultimate aims, tools utilising population scale information can guide not only current efforts towards a conclusion, but also generate new avenues of study. We leave the above with a great many questions to be answered in both bodies of research. However, the findings garnered here provide guides into new avenues of study. Progress continues in both research questions, and it is hoped that the current document may provide some guide to this new body of study towards reaching final conclusions.

6.1 References

Andrews, G.P. et al., 1996. Fraction 1 capsular antigen (F1) purification from Yersinia pestis CO92 and from an Escherichia coli recombinant strain and efficacy against lethal plague challenge. *Infection and Immunity*, 64(6), pp.2180–2187.

Baker, E.E. et al., 1952. Studies on Immunization Against Plague: I. The Isolation and Characterization of the Soluble Antigen of Pasteurella Pestis. *J Immunol*, 68, pp.131–145.

Basáñez, M.G., Churcher, T.S. & Grillet, M.E., 2009. Chapter 11 Onchocerca-Simulium Interactions and the Population and Evolutionary Biology of Onchocerca volvulus. In *Advances in Parasitology*. pp. 263–313.

Begon, M. et al., 2006. Epizootiologic parameters for plague in Kazakhstan. *Emerging infectious diseases*, 12(2), pp.268–73.

Colebunders, R. et al., 2016. Risk factors for epilepsy in Bas-Uélé Province, Democratic Republic of the Congo: A case-control study. *International Journal of Infectious Diseases*, 49, pp.1–8.

Davis, K.J. et al., 1996. Pathology of experimental pneumonic plague produced by Fraction 1-Positive and Fraction 1-Negative Yersinia Pestis in African Green Monkeys (Cercopithecus aethiops). *Archives of pathology & laboratory medicine*, 120(February).

Davis, S., et al., 2007a. Empirical assessment of a threshold model for sylvatic plague. *Journal of the Royal Society, Interface / the Royal Society*, 4(15), pp.649–57.

Davis, S., et al., 2007b. Plague metapopulation dynamics in a natural reservoir: the

burrow system as the unit of study. *Epidemiology and infection*, 135(5), pp.740–8.

Davis, S. et al., 2004. Predictive thresholds for plague in Kazakhstan. *Science (New York, N.Y.)*, 304(5671), pp.736–8.

Derbise, A. et al., 2015. Complete Protection against Pneumonic and Bubonic Plague after a Single Oral Vaccination. , pp.1–19.

Dowell, S.F. et al., 2013. Nodding syndrome. *Emerging Infectious Diseases*, 19(9), pp.1374–1384.

Drozdov, I.G. et al., 1995. Virulent non-capsulate Yersinia pestis variants constructed by insertion mutagenesis. *Journal of Medical Microbiology*, 42(4), pp.264–268.

Druet-Cabanac, M. et al., 1999. Onchocerciasis and epilepsy: a matched case-control study in the Central African Republic. *American journal of epidemiology*, 149(6), pp.565–570.

Galen, J.E. et al., 2015. A Bivalent Typhoid Live Vector Vaccine Expressing both Chromosome- and Plasmid-Encoded Yersinia pestis Antigens Fully Protects against Murine Lethal Pulmonary Plague Infection. *Infection and Immunity*, 83(1), pp.161–172.

Gibbs, E.P.J., 2014. The evolution of One Health: a decade of progress and challenges for the future. *The Veterinary record*, 174(4), pp.85–91.

Hanski, I. & Ovaskainen, O., 2003. Metapopulation theory for fragmented landscapes. *Theoretical Population Biology*, 64(1), pp.119–127.

Idro, R. et al., 2016. Is nodding syndrome an Onchocerca volvulus-induced neuroinflammatory disorder? Uganda's story of research in understanding the disease. *International Journal of Infectious Diseases*, 45, pp.112–117.

Jesse, M. & Heesterbeek, H., 2011. Divide and conquer? Persistence of infectious agents in spatial metapopulations of hosts. *Journal of Theoretical Biology*, 275(1), pp.12–20.

Johnson, T. et al., 2014. Detection of auto-antibodies to leiomodin-1 in patients with nodding syndrome. *Journal of Neuroimmunology*, 275(1–2), p.103.

Jones, K.E. et al., 2008. Global trends in emerging infectious diseases. *Nature*, 451(7181), pp.990–3.

Kaiser, C., Pion, S.D.S. & Boussinesq, M., 2013. Case-control Studies on the Relationship between Onchocerciasis and Epilepsy: Systematic Review and Meta-analysis. *PLoS Neglected Tropical Diseases*, 7(3).

Kausrud, K.L. et al., 2007. Climatically driven synchrony of gerbil populations allows large-scale plague outbreaks. *Proceedings of the Royal Society: Biological Sciences*, 274(1621), pp.1963–9.

König, R. et al., 2010. The role of Onchocerca volvulus in the development of epilepsy in a rural area of Tanzania. *Parasitology*, 137(10), pp.1559–68.

Laperrière, V., Brugger, K. & Rubel, F., 2016. Cross-scale modeling of a vector-borne disease, from the individual to the metapopulation: The seasonal dynamics of

sylvatic plague in Kazakhstan. *Ecological Modelling*, 342, pp.34–48.

Levick, B. et al., 2015. The perfect burrow, but for what? Identifying local habitat conditions promoting the presence of the host and vector species in the Kazakh plague system. *PLoS ONE*, 10(9), p.e0136962.

Moran, N.A., McCutcheon, J.P. & Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. *Ann Rev Genet*, 42(1), pp.165–190.

Naumov, N.P. & Lobachev, V.S., 1975. Ecology of desert rodents of the USSR (Jerboas & Gerbils). In I. Prakash & P. K. Ghosh, eds. *Rodents in Desert Environments*. The Hague: The Hague, pp. 465–598.

Newell, E.D., Vyungimana, F. & Bradley, J.E., 1997. Epilepsy, retarded growth and onchocerciasis, in two areas of different endemicity of onchocerciasis in Burundi. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 91(5), pp.525–527.

Pion, S.D.S. et al., 2009. Epilepsy in Onchocerciasis Endemic Areas: Systematic Review and Meta-analysis of Population-Based Surveys. *PLOS Neglected Tropical Diseases*, 3(6), p.e461.

Reijniers, J. et al., 2014. Plague epizootic cycles in Central Asia. *Biology letters*, 10(6), pp.0–3.

Salathe, M. & Jones, J.H., 2010. Dynamics and Control of Diseases in Networks with Community Structure. *PLoS computational biology*, 6(4), p.e1000736.

Schmid, B. V. et al., 2012. Local persistence and extinction of plague in a metapopulation of great gerbil burrows, Kazakhstan. *Epidemics*, 4(4), pp.211–218.

Shenbrot, G. et al., 2002. Habitat-dependent differences in architecture and microclimate of the burrows of Sundevall's jird (Meriones crassus) (Rodentia: Gerbillinae) in the Negev Desert, Israel. *Journal of Arid Environments*, 51(2), pp.265–279.

Simonsen, P.E., 2009. Filariases. In G. C. Cook & A. I. Zumla, eds. *Manson's Tropical Diseases*. pp. 1491–1500.

Simpson, W.J., Thomas, R.E.X.E. & Schwan, T.O.M., 1990. Recombinant capsular antigen (Fraction 1) from Yersinia pestis induces a protective antibody response in BALB/C Mice. *Americal Journal of Tropical Medicine and Hygiene*, 43(4), pp.389–396.

Titball, R.W. et al., 1997. Expression of the Yersinia pestis capsular antigen (F1 antigen) on the surface of an aroA mutant of Salmonella typhimurium induces high levels of protection against plague. *Infection and Immunity*, 65(5), pp.1926–1930.

Wamala, J.F. et al., 2015. Nodding Syndrome May Be Only the Ears of the Hippo. *PLOS Neglected Tropical Diseases*, 9(8), p.e0003880.

Weening, E.H. et al., 2011. The dependence of the Yersinia pestis capsule on pathogenesis is influenced by the mouse background. *Infection and Immunity*, 79(2), pp.644–652.

Whitford, W.G. & Kay, F.R., 1999. Biopedturbation by mammals in deserts: a review. *Journal of Arid Environments*, 41(2), pp.203–230.

WHO, 2013. WHO Global Health Data Repository. Available at: http://apps.who.int/gho/data/view.main.1540_52?lang=en.

Wilschut, L.I. et al., 2013. Mapping the Distribution of the main host for plague in a complex lanscape in Kazakhstan : An object-based approach using SPOT-5 XS , Landsat 7 ETM + , SRTM and multiple Random Forests. *International Journal of Applied Earth Observation and Geoinformation*, 23, pp.81–94.

Wilschut, L.I. et al., 2015. Spatial distribution patterns of plague hosts: point pattern analysis of the burrows of great gerbils in Kazakhstan. *Journal of Biogeography*, 42(7), pp.1281–1292.

Xu, W. et al., 2012. Rhombomys opimus contribution to the “fertile island” effect of tamarisk mounds in Junggar Basin. *Ecological Research*, 27(4), pp.775–781.

Zouré, H.G.M. et al., 2014. The geographic distribution of onchocerciasis in the 20 participating countries of the African Programme for Onchocerciasis Control : (2) pre-control endemicity levels and estimated number infected. *Parasites & vectors*, pp.1–15.

Appendices

Contents

Table I: Details of studies on the virulence of the F1 capsular antigen.

Study	Host species	Strain/genotype	Method of infection
Burrows 1957	Mice, guinea pigs.	Y.pestis M23 F1-.	
Donavan 1961	Guinea Pigs	Y. pestis M23 F1-.	Intradermal inoculation.
Drozdov 1995	Mice, guinea pigs.	Y.pestis transformed with pFS23 plasmid.	Subcutaneous injection.
Friedlander 1995	Swiss Webster mice, Hartley guinea pigs.	Y.pestis K25 (66bp deletion in caf1)	Aerosol
		Java 9 (cryptic 37kb pFra plasmid).	Subcutaneous injection (mice only).
Samoilova 1996	Guinea pigs.	Y.pestis 358 F1- forms.	Aerosol
			Subcutaneous injection.
Davis 1996	Green monkeys.	Y.pestis F1- CO92-C12.	Aerosol.
Queenee 2008	BALB/c mice.	Y.pestis KIMD27 (pgm-), compared to non-attenuated CO92.	Subcutaneous injection.
Cornelius 2009	BALB/c mice.	Y.pestis Caf1A mutant.	Intramuscular injection.
Weening 2011	BALB/c J mice,	Y.pestis CO92 double mutants.	Intraperitoneal injection.
	C57BL/6J mice.		

Table II: studies of the immunogenicity of the F1 capsular antigen.

Study	Host	Vaccination	Challenge
Baker 1952	ABC mice, white rats, guinea pigs	Subcutaneous injection of purified F1 antigen from Yreka Y.pestis	Subcutaneous injection of Shasta Y.pestis
Simpson 1990	BALB/c mice	Subcutaneous injection of purified F1 antigen produced by transformed E.coli or whole cell E.coli producing F1.	Intraperitoneal injection of Y.pestis.
Andrews 1996	Swiss Webster mice	Subcutaneous injection of purified F1 antigen from CO92 Y.pestis and whole cell vaccine.	Aerosol and Intravenous injection of CO92 Y.pestis.
Titball 1997	BALB/c mice	Subcutaneous injection of F1 antigen or whole cell S.typhimurium expressing F1.	Subcutanous injection of GB Y.pestis.
Batra 2014	BALB/c mice	Subcutaneous injection with purified F1 protein or combination with one/both of HSP70(II) or LcrV.	Intraperitoneal injection of S1 Y.pestis.
Derbise 2015	OF1 or C57BL/6 mice	Intragastric administration of whole cell attenuated V674 Y.pseudotuberculosis carrying caf operon.	Subcutanous injection of CO92 Y.pestis.
Galen 2015	BALB/c mice	Intransal immunisation with live S.typhi expressing either one or both of F1 and LcrV, boosted with intramuscular injection of pure LcrV.	Intranasal challenge with CO92 Y.pestis.

Binomial logistic regression model results and selection: association of latitude location (°N), longitude location (°E) and year with plague.

Table I: Full model including all factors.

	Coefficient	Standard Error	P Value
Intercept	149248.16	25032.81	<0.0001
Latitude	-444.69	357.95	0.21
Longitude	-1698.79	235.54	<0.0001
Year	-74.29	12.46	<0.0001
Latitude:Year	0.22	0.18	0.21
Longitude:Year	0.85	0.12	<0.0001

Table II: Single term deletions. A non-significant improvement in AIC is achieved by removing Latitude:Year, so this is removed to increase model simplicity.

Single term deletions				
	DF	Deviance	AIC	ΔAIC
Latitude:Year	1.00	505.71	515.71	-0.41
Longitude:Year	1.00	566.44	567.44	51.32

Table III: Model with latitude:year removed.

	Coefficient	Standard Error	P Value
Intercept	128205.77	18073.39	<0.001
Latitude	1.59	0.37	<0.001
Longitude	-1684.055	237.02	<0.001
Year	-63.81	8.99	<0.001
Longitude:Year	0.84	0.12	<0.001

Table IV: Single term deletions. Removing either term would significantly diminish AIC, so both are retained to obtain a minimum model.

	DF	Deviance	AIC	ΔAIC
Lat	1.00	527.09	535.09	19.38
Longitude:Year	1.00	568.39	576.39	60.68

Binomial logistic regression model results and selection: association of latitude location (°N), longitude location (°E) and year with $F1^-$ isolate presence.

Table I: Full model results.

	Coefficient	Standard Error	P Value
Intercept	13466.44	17878.03	0.45
Latitude	-74.93	219.77	0.73
Longitude	-133.65	126.09	0.29
Year	-6.70	8.90	0.45
Latitude:Year	0.04	0.11	0.73
Longitude:Year	0.07	0.06	0.29

Table II: Single term deletions. Removing neither term creates significant improvements in AIC, so the full model is maintained.

	DF	Deviance	AIC	ΔAIC
Latitude:Year	1.00	1015.90	1025.90	-1.80
Longitude:Year	1.00	1016.90	1026.90	-0.80

Chi squared test results of $F1^-$ isolate presence between rodent demographics.

Table I: Sex and F1 status of strain isolated.

	F1+	F1-	N
Female	0.447368	0.552632	76
Male	0.516484	0.483516	91
X_2	0.54		
P	0.46		

Table II: Age (adult or juvenile) and F1 status of strain isolated.

	F1+	F1-	N
Adult	0.433333	0.566667	90
Juvenile	0.545455	0.454545	77
X_2	1.66		
P	0.2		

Table III: Reproductive status in adult females and F1 status of strain isolated.

	F1+	F1-	N
Adult female	0.380952	0.619048	21
Adult female (pregnant or lactating	0.32	0.68	25
X_2	0.01		
P	0.9		

Nearest burrows with plague where burrows transition from no *Yersinia* to F1$^-$ *Yersinia.*

In 2 instances, (at burrow 10 in 2007 and at burrow 31 in 2007), F1- plague is seen where the only previous observation was of plague absence. In the years between the last observation at the burrow, and the observation of an F1- strain, the nearest burrows in each year (within 1km of the burrow) with F1+ plague were identified

Table I: Burrow 10, found to have F1$^-$ *Yersinia* in 2007, following the last observation of no *Yersinia* in 2003.

Year	Nearest Burrow	Distance between burrow and burrow 10 (km)
2003	11	0.007
2004	9	0.005
2005	30	0.005

Table II: Burrow 31, found to have F1$^-$ *Yersinia* in 2007, following the last observation of no *Yersinia* in 2003.

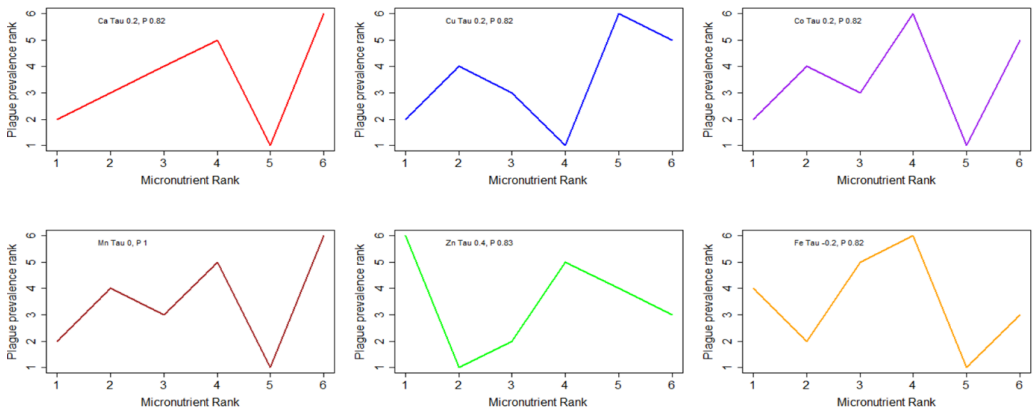
Year	Nearest Burrow	Distance between burrow and burrow 10 (km)
2003	53	0.0007
2004	9	0.08
2005	12	0.08

Kendall's Rank test of correlation results: sector rank by trace element concentration and number of *Yersinia* isolates.

Table I: Results of Kendall's Rank tests.

Data 1	Data 2	Tau	P
Average calcium concentration (2011-2012)	Average *Yersinia* prevalence (2000-2013)	-0.2	0.72
Average copper concentration (2011-2012)	Average *Yersinia* prevalence (2000-2013)	-0.07	1
Average cobalt concentration (2011-2012)	Average *Yersinia* prevalence (2000-2013)	0.07	1
Average iron concentration (2011-2012)	Average *Yersinia* prevalence (2000-2013)	-0.2	-0.72
Average manganese concentration (2011-2012)	Average *Yersinia* prevalence (2000-2013)	-0.33	-0.47
Average zinc concentration (2011-2012)	Average *Yersinia* prevalence (2000-2013)	0.6	-0.14
Average Fe:Cu ratio (2011-2012)	Average *Yersinia* prevalence (2000-2013)	-0.07	1
Average Fe:Co ratio (2011-2012)	Average *Yersinia* prevalence (2000-2013)	-0.2	0.72

Figure I: Sector ranks for each micronutrient and *Yersinia*.

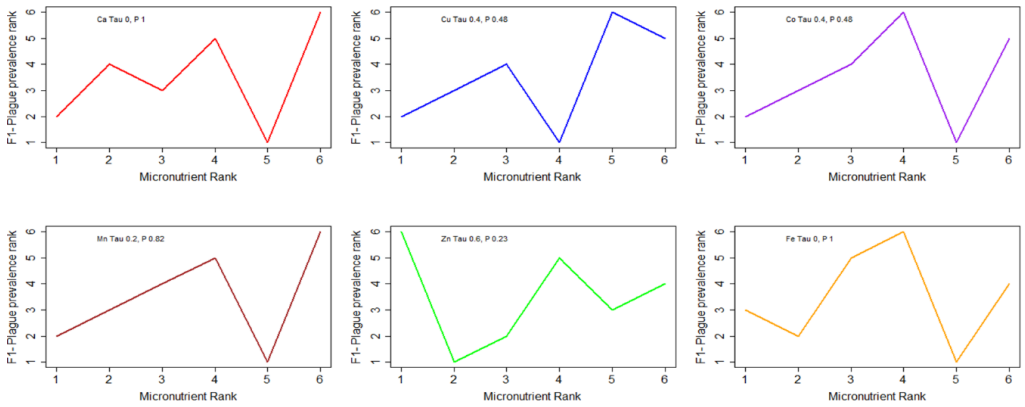


Kendall's Rank test of correlation results: sector rank by trace element concentration and number of F1$^-$ *Yersinia* isolates.

Table I: Results of Kendall's Rank tests.

Data 1	Data 2	Tau	P
Average calcium concentration (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	0.33	0.47
Average copper concentration (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	0.47	0.27
Average cobalt concentration (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	0.33	0.47
Average iron concentration (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	0.07	1
Average manganese concentration (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	0.47	0.27
Average zinc concentration (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	0.07	1
Average Fe:Cu ratio (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	0.2	0.72
Average Fe:Co ratio (2011-2012)	Average F1- *Yersinia* prevalence (2000-2013)	-0.47	0.27

Figure I: Sector ranks for each micronutrient and F1$^-$ *Yersinia*.



Cross Correlation analysis of time series.

Figure I: Gerbil occupancy of burrows and proportion of F1$^-$ *Yersinia* isolates, 2005-2013.

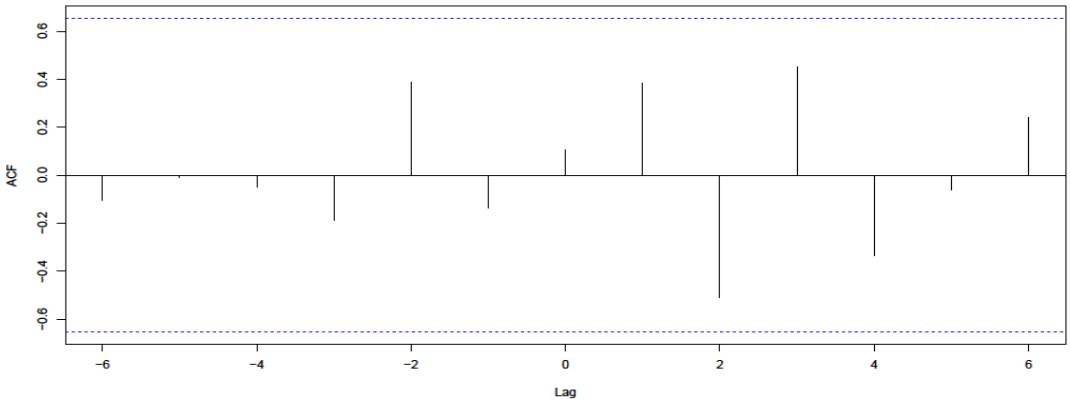
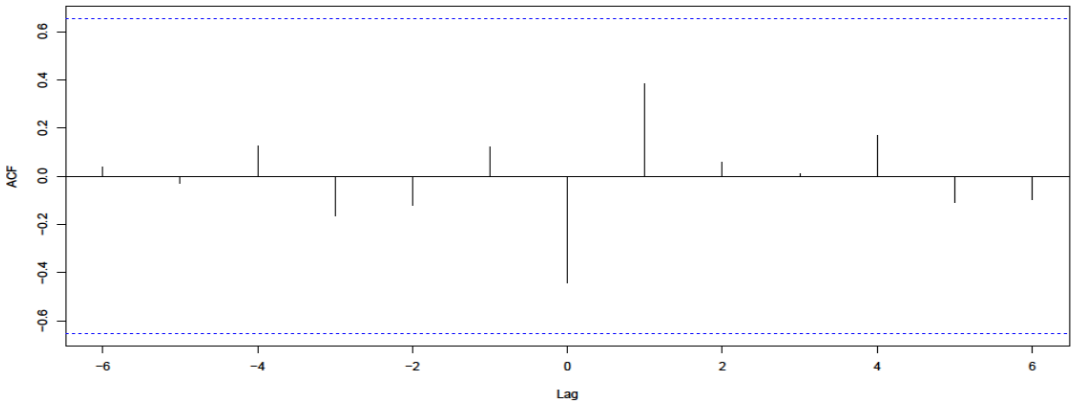


Figure II: *Yersinia* isolates, and the proportion of those isolates that were F1$^-$.



Field Methods

The data described here were collected by Kazakh public health teams in 1995, and from 1999-2013 in the desert region south of Lake Balkhash in the South-Eastern of Kazakhstan. Field data include details of specific animals trapped: both *R. opimus* individuals (for whom life stage and sex were recorded) and insect vectors for individual burrows.

Alongside normal recording of *Y. pestis* presence across the Pre-Balkhash region, records were also generated detailing the F1 phenotype of isolates that were recovered from the infected hosts and ectoparasites (fleas and ticks). *Y. pestis* strains were isolated from infected rodent hosts by plating blood and tissue samples on Hottinger's agar or from ectoparasites by crushing them in saline and plating on the same agar. *Y. pestis* identity was confirmed in the laboratory by examining morphology, growing speed, sugar fermentation and urea metabolism. Samples were confirmed as *Y. pseudotuberculosis* negative using the bacteriophage L-413C (Fillipov et al., 2012). Strains' ability to produce the F1 pseudocapsule was determined by double diffusion in gel, passive haemagglutination (PHA), and enzyme linked immune absorbent assay (ELISA) detecting antibody response (Bektemirov et al., 1986; Martinevsky, 1973).

Records estimating overall gerbil occupancy levels were also collected by the public health teams in a separate study. These are recorded as the number of gerbils observed per sector (10 × 10 km) area, observed once in the spring season (May-August) and once in the autumn season (September-November) of each year.

Field accounts from natural plague foci in Uzbekistan and China have identified correlations between local levels of trace elements including iron, copper, cobalt and zinc in the environment and the location of epizootics in the foci (Gage and

Kosoy, 2005; Liu, 2000; Mezentsev et al., 2000a). Similar observations have been made in plague positive areas of Tanzania (Meliyo et al., 2015). Additionally, laboratory studies have identified differential progression of *Y. pestis* infection in great gerbils subjected to diets with differential micronutrient profiles (Mezentsev et al., 2000b). With a view to identifying whether similar relationships exist in the Pre-Balkhash plague focus, perhaps related to the F1 phenotype, concentrations of six trace elements (zinc, cobalt, copper, iron, calcium, and manganese) were recorded at individual burrow locations across 6 study sectors in 2011 and 2012. These elements are known to regulate the expression of several *Y. pestis* genes (including virulence genes) (Bearden and Perry, 1999; Carniel et al., 1987; Hu et al., 1998), and have been shown to determine host response to infection and infection outcomes across a range of mammal species, including rodents (Kubena and McMurray, 1996). Specifically, the trace element composition at *R. opimus* burrow sites was recorded using plant stomach contents of the hosts, as these are representative of the true food intake, combining the hosts' food plants and any mixed soil consumed by the hosts.

In order to standardise the measurements, and to avoid age and sex related variation, the stomach content of 5 adult males were sampled per square (500 m × 500 m or 200 m × 200 m in Spring 2013, see (Levick et. al, 2015)).When the number of individuals was too low, sub-adult males were selected (32 out of 376 samples in total) and then ultimately females were taken when no males had been caught (62 samples). Stomach contents were stored in vials, sterilized (121°C, 15 min, 2.2 bar) and dried in the Bakanas anti-plague station laboratory after necropsy of the rodent. The concentration of the elements was measured for each sample, including blanks of 0.5 ml water and duplicates to act as internal controls. This was performed using atomic absorption spectrophotometry by SAC (Almaty, Kazakhstan) according

to certified procedures (CHEMEX) and calibrated equipment (serial dilutions of FLUKA solutions in bi-distilled water or 0.1 $MHNO_3$ depending on the element). The analysis was conducted three times in each year (spring, summer and autumn), and in order to estimate the precision and the reproducibility of the results, certified reference material (CRM; soil samples with known concentrations of different elements) were included in each round, and the method detection limits (MDL) were determined for each analysed element for each round.

The analyses were performed on 0.5 g of stomach content with concentrations expressed as milligrams per kilogram dry weight. Samples were gritted,dried at 100°C and ashed at 450-500°C in the presence of ammonium sulphate for 3 hours. Ashed residues were diluted with nitric acid before being submitted to atomic absorption spectrometry with flame atomization (atomic absorption spectrophotometer AAS 8010, Younglin, South Korea). When necessary the deuterium corrector was used as mentioned in the international standard (GOST 30178-96). The wavelengths of the reads were set at 422.7 nm for Ca, 324.8 nm for Cu, 213.9 nm for Zn, 248.3 nm for Fe, 403.1 nm for Mn and 240.7 nm for Co.